Unravelling (Mis)conceptions about Algebraic Letters: Exploring Response Patterns in the 'Meaning of Letters' SMART-test using Latent Class Analysis

Katrin Klingbeil[1] and Filip Moons[2]

[1]University of Duisburg-Essen, Germany; katrin.klingbeil@uni-due.de

[2]Freudenthal Institute, Utrecht University, The Netherlands; f.moons@uu.nl

Identifying typical hurdles, common errors and misconceptions in a certain domain is crucial to deepen our understanding of students' learning. In this paper, we explore response patterns of 2051 German grade 7 and 8 students shown in the SMART-test "Meaning of Letters", designed to assess (mis)conceptions regarding variables, more precisely, the letter-as-object misconception. Using Latent Class Analysis, we were able to identify six response patterns. These patterns are described and thoroughly analysed. They urge us to think more deeply about the interplay between (mis)conceptions and contexts and can help build valid assessment tools to diagnose students' current understanding.

Keywords: Online formative assessment, algebra, variables, letter-as-object misconception, latent class analysis

Introduction

Assessing students' (mis)conceptions is a challenging task. SMART ("Specific Mathematics Assessments that Reveal Thinking") online tests, that have been developed at the University of Melbourne since 2008, offer a solution by facilitating easy provision and processing of diagnostic tasks on students' conceptual understanding and potential misconceptions. SMART's extended analysis detects patterns between diagnostic tasks, revealing insights into students' understanding and misconceptions. In addition to this automatic diagnosis, it also provides teachers with explanations, tasks, and suggestions for targeted interventions (Steinle et al., 2009).

The investigated test here, *Meaning of Letters*, aims to assess the *letter-as-object* misconception and its subtypes based on students' responses to six multiple-choice tasks. Despite known challenges of multiple-choice tasks, developers argue that well-designed task can effectively unveil students' thinking: Klingbeil et al. (2024) showed that students' explanations aligned well with their shown (mis)understandings in their multiple-choice responses.

In a comprehensive intervention study spanning six federal states in Germany, 2051 7th- and 8th-grade students undertook the *Meaning of Letters* test after a few algebra lessons. This paper investigates the response patterns of these students using *Latent Class Analysis* (LCA) (Brandenburger & Schwichow, 2023). LCA is a statistical method used to identify unobservable subgroups (latent classes) within a heterogeneous population based on patterns of responses. This analysis can unravel the (mis)conceptions in understanding algebraic letters and how they interact.

In the following paragraphs, we introduce the theoretical background behind the *Meaning of Letters*-test and the six tasks; next we pose the research question.

Struggling to understand algebraic letters: the *letter-as-object* misconception and its subtypes

Arcavi, Drijvers and Stacey (2017) "distinguish five facets of the concept of variable: a placeholder for a number, an unknown number, a varying quantity, a generalised number, and a parameter" (p. 12). Across these facets, variables stand for or refer to one or more numerical values. Yet, algebra learners often struggle with this numerical interpretation, and various typical errors and misconceptions have been identified. One of them, the *letter-as-object* (LO) misconception, has been described by Küchemann in 1981 as the letter being "regarded as a shorthand for an object or as an object in its own right" (p. 104) and extensively documented over decades (e.g., Akhtar & Steinle, 2017). As part of the foundational Concepts in Secondary Mathematics and Science (CSMS) study on the mathematical understanding of secondary school students in the United Kingdom, Küchemann (1981) utilised the following task: "Blue pencils cost 5 pence each, and red pencils cost 6 pence each. I buy some blue and some red pencils and altogether it costs me 90 pence. If b is the number of blue pencils bought and if r is the number of red pencils bought, what can you write down about b and r?" (p. 107).

While only 10% of tested 14-year-old students provided the correct equation $5b + 6r = 90$, 17% gave $b + r = 90$ as an answer, which might have been read as "blue pencils and red pencils together cost 90 pence" (LO). Interpreting b as "the number of blue pencils" is a possibility here, too; however, this would still imply a wrong understanding of equations with a number of pencils on one side of the equation and the price of all pencils on the other. Interestingly, 6% of the students came up with another kind of equation: $6b + 10r = 90$ or $12b + 5r = 90$. These students had figured out a possible solution to the problem first and then used these values as coefficients in their equation. Since the letters are used as abbreviations for the involved objects ("12 blue pencils and 5 red pencils together cost 90 pence"), this is regarded as a special form of LO, which we will refer to as the s*olution-as-coefficient* (SAC) misconception in the following. Another special form of LO is called *letter-as-unit* (LU) when the algebraic letter is interpreted as an abbreviation for a unit (Akhtar & Steinle, 2017), e.g., in a task about 8 trucks weighing 24 tonnes, the t in the equation $8t = 24$ would be misinterpreted as standing for tonnes (not realising that this would not be a correct equality).

Research question

The *Meaning of Letters* SMART-test polls the understanding of variables and detects the presence of the letter-as-object misconception, leading to the following research question: *Which response patterns regarding the letter-as-object misconception can be identified among German grade 7 and 8 students based on their responses to the six multiple-choice tasks of the SMART-test Meaning of Letters?*

Methods and Materials

SMART-test *Meaning of Letters*

For the diagnosis of students, two parallel versions of the SMART-test *Meaning of Letters* were used with the A or B version randomised by class. Here, we describe only the A version of the German translation of the test (see Figure 1; COR indicating the correct response option). The first task type (Meaning tasks), originating from the work of MacGregor and Stacey (1997), uses only one algebraic letter and asks students to decide on the meaning of this letter in a linear equation in a given context. While the *Ducks* item uses the initial letter of the involved objects and units, the *Bricks* item uses the letter y. Apart from the correct response (cost/height), MC options include the involved objects (singular and plural; LO) as well as the corresponding unit (LU).

Meaning tasks	**Additive tasks**	**Proportional tasks**
Ducks	*Garden*	*Biros*
Lucy bought 6 ducks for \$12. She wrote the equation $6d = 12$. What does d in Lucy's equation stand for? ☐ the cost of one duck (COR) ☐ one duck (LO) ☐ ducks (LO) ☐ dollars (LU)	Payam bought r red rose bushes and l lilac lavender bushes. The roses cost €4 each. The lavenders cost €5 each. Which equation says that the total cost was \$70? ☐ $4r + 5l = 70$ (COR) ☐ $r + l = 70$ (LO) ☐ $10r + 6l = 70$ (SAC)	Biros are sold in packs of 3. Sam bought p packs and got b biros altogether. Choose the correct equation: ☐ $3p = b$ (COR) ☐ $p = 3b$ (LO) ☐ $p = 3$ (LO) ☐ $b + p = 4$ (LO) ☐ $30b = 10p$ (SAC)
Bricks	*Wheels*	*Racetrack*
Tina stacked 9 identically sized bricks on top of each other making a 99 mm high tower. She wrote the equation $9y = 99$. What does y in Tina's equation stand for? ☐ the height of one brick (COR) ☐ one brick (LO) ☐ the bricks in the tower (LO) ☐ millimetres (LU)	At a bike shop there are b bikes (2 wheels) and t trikes (3 wheels). Which equation says that there is a total of 100 wheels in the shop? ☐ $2b + 3t = 100$ (COR) ☐ $b + t = 100$ (LO) ☐ $35b + 10t = 100$ (SAC)	A car takes 12 minutes to drive round this racetrack. A driver drives r times around the racetrack in m minutes. Choose the correct equation: ☐ $12r = m$ (COR) ☐ $12m = r$ (LO) ☐ $r = 12$ (LO) ☐ $5r = 60m$ (SAC)

Figure 1: Tasks of Meaning of Letters test (German A version) translated back into English

The second type of task (Additive tasks) is based on Küchemann (1981). It uses two algebraic letters (corresponding to the initial letters of involved objects), which are additively connected and restricted by the given situation. Students are supposed to choose the correct linear equation (in standard form) for the described context. In the correct equation, the letters represent the number of objects and the coefficients for the price per object (*Garden*) and the number of components per object (*Wheels*), respectively. The first alternative response option simply adds the variables without any coefficients, making it possible to interpret the letters as abbreviations for the involved objects (e.g., "Bikes and trikes have 100 wheels altogether."; LO). In the equation of the other alternative option, coefficients equal a possible solution to the problem (that has not been posed) so that the equation can be read as some solution sentence (e.g., "35 bikes (with 2 wheels each) plus 10 trikes (with 3 wheels each) have 100 wheels altogether."; SAC). Also, in this case, the letters are read as abbreviations for the involved objects.

The third task type (Proportional tasks) is derived from the famous "Students and Professors" problem (Clement et al., 1981):

> *"There are six times as many students as professors at this university." Write an equation using S for the number of students and P for the number of professors.*

This is often answered with $6S = P$ instead of $6P = S$. The proportional tasks in the SMART-test have the same algebraic structure: the two variables are directly proportional to each other. Students are again asked to choose the equation matching the given situation. In the correct equation, the letters (matching the initial letters of involved objects/units) stand for the number of objects for both involved objects (*Biros*) or for the number of racetrack rounds and the number of minutes (*Racetrack*).

In both items, the coefficient is the proportionality constant (number of biros per pack or number of minutes per round). The first alternative response option is the reverse of the correct equation, which allows for a LO interpretation (e.g., “A pack contains 3 biros.” or “1 round equals 12 minutes”; LO). For the *Racetrack* task, the first alternative can also be seen as an LU interpretation (e.g., “12 minutes equals one round). However, since it is unclear how exactly students interpret the letter here, we opted for the more general LO interpretation. The LO interpretation also applies to the second alternative response although the second variable is missing (e.g., as “A pack has 3.”; LO). In the equation of the third alternative option, the coefficients correspond to a possible solution (to the question that has not been asked), which can be interpreted as a kind of solution sentence (e.g., “Sam bought 10 packs and has 30 biros now.”), indicating the SAC misconception. The *Biros* task offers one more response option that features the addition of the two variables without coefficients. Again, the letters can be interpreted as abbreviations (e.g., as “One biro plus a pack of biros is 4 altogether.”; LO). This response type does not make sense for the *Racetrack* task since no different objects but rounds and minutes would be added.

Participants

In total, 2051 grade 7 and 8 students (aged 12–14) from six federal states of Germany (78% attending grammar schools, 22% attending non-grammar schools) completed the SMART test. These students were taught by 103 mathematics teachers, leading to a nested data structure (data are analysed at the student level, but the student are clustered in classrooms).

Teachers were asked to administer the SMART online test 1–2 weeks into their teaching sequence about variables, algebraic expressions and/or equations. Thus, students in grade 7 should have been familiar with a basic concept of variables and be able to use and manipulate them in easy algebraic expressions before taking the test. In grade 8, students probably have started focussing on (solving) equations.

Data analysis

The response patterns of the test were analysed using Latent Class Analysis (LCA) (Brandenburger, & Schwichow, 2023). LCA is a form of structural equation modelling useful for identifying patterns/groups within categorical responses. These patterns/groups are called *latent classes.* Intuitively, one can think of the 2051 participating students as ‘latent classes’. Of course, a description of 2051 ‘latent classes’ will be hard to interpret. LCA considerably reduces the complexity of the data by grouping students with similar patterns of responses in one class, bringing down the 2051 ‘latent classes’ to a comprehensible, clearly distinguishable number of latent classes.

We used SAS Enterprise Guide 8.3 with the PROC LCA for the LCA analysis, considering the nested data structure. As students were not required to answer all questions, 49 students of the 2051 (2.3%) left some tasks unanswered. Hot-deck imputation was used to impute these missing values.

Results

Overall response rates to the SMART-test *Meaning of Letters*

The overall response rates of the participants are shown in Figure 2. Note the low number of correct answers as well as that LU answers are only possible in task 1 and 2, while SAC answers are only possible in tasks 3 to 6. The LO misconception was omnipresent in the responses to the meaning and proportional tasks; while SAC was present in most responses to the additive tasks.

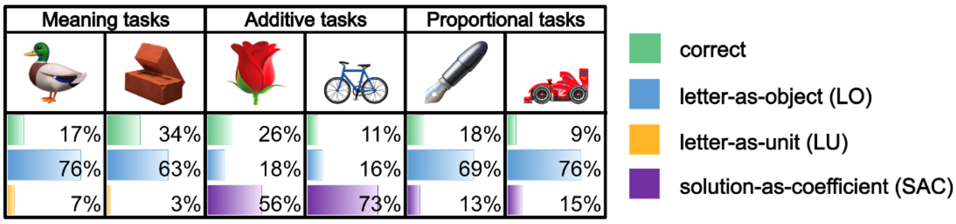


Figure 2: Overall response rates to the *Meaning of Letters* test

Model selection and model fit of LCA

We iteratively built several models with different numbers of classes to choose the appropriate number of classes for our LCA model. Information criteria (AIC/BIC) and the possibility of giving meaningful labels to the classes were used to decide the number of classes. The 6-class model had the best information criteria, the most interpretable latent classes, and no particularly infrequent class. Table 1 presents the given labels to each class and the prevalence (indicating overall class membership probability) of each class (last row). There is no order or hierarchy of the classes.

			Mean classification probability					
Classes with labels		Preva-lence	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Class 1	LO predominant	23%	**85%**	2%	5%	1%	6%	1%
Class 2	LO with SAC	11%	4%	**91%**	1%	3%	2%	0%
Class 3	LO with SAC only in additive tasks	38%	6%	1%	**91%**	2%	1%	0%
Class 4	Correct meaning tasks with LO/SAC elsewhere	8%	4%	3%	9%	**77%**	3%	3%
Class 5	LO apart from additive tasks	15%	8%	0%	15%	2%	**70%**	4%
Class 6	Mostly correct	5%	3%	1%	1%	6%	8%	**81%**

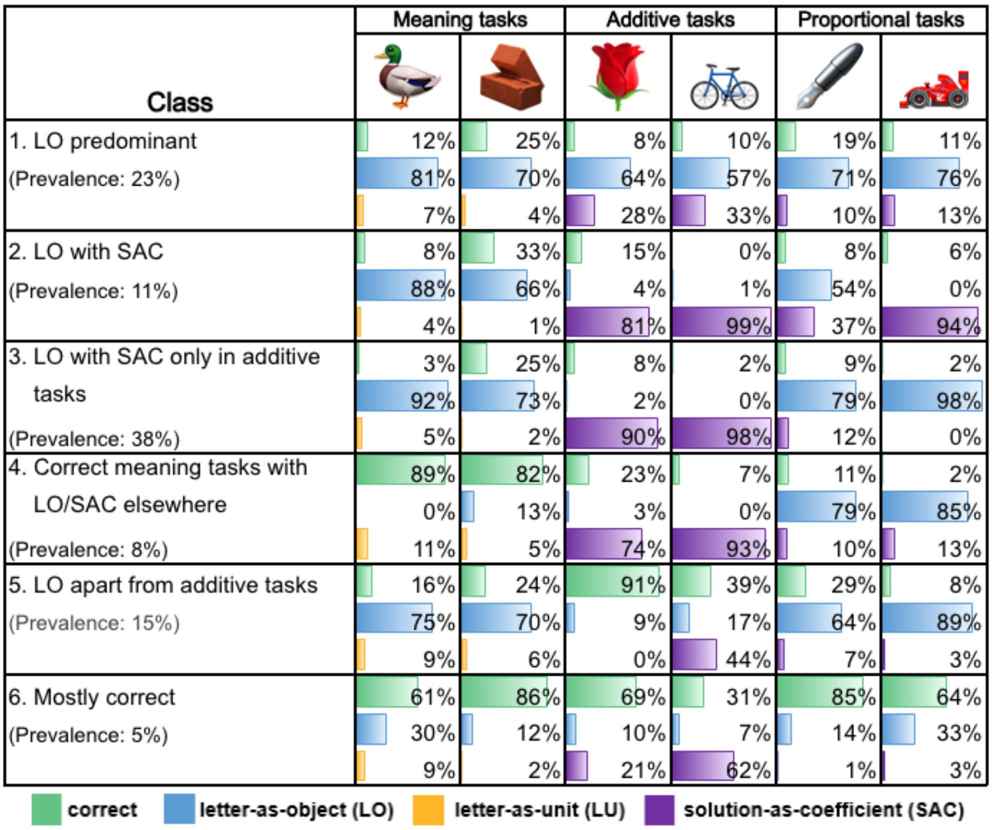
Table 1: Mean classification probability; hit rate in the diagonal (bold)

LCA allows the calculation of the likelihood that a student belongs to each class by analysing their test responses. A robust LCA model, characterised by high homogeneity among latent classes and clear class separation, yields for most students a class where the probability of classification is high for that best-fitting class and low for the others. By calculating the mean classification probability for all students aggregated along their best-fitting class, we get a grip on the homogeneity and class separation in our 6-class LCA model. These mean classification probabilities are shown in Table 1: for example, a student with class 2 as best has a probability of 91% to be in this class (we call this the 'hit rate', shown in the diagonal of the table) and only a probability of 4% to be in the first.

Description of the latent classes

In Figure 3, the item-response probabilities for each task are shown for the six classes. In the following paragraphs, we give a description of each of the classes.

Class 1 (prevalence: 23%) is characterised by LO responses being most likely in all tasks. Even when the initial letter of the involved object is not used (*Bricks*), LO is more likely than a correct answer. The subtype SAC is also possible in additive tasks, but less likely than LO, indicating a relatively consistent interpretation of letters as abbreviations for objects. We label this class "*LO predominant"*



Class	Meaning tasks		Additive tasks		Proportional tasks	
1. LO predominant	12%	25%	8%	10%	19%	11%
(Prevalence: 23%)	81%	70%	64%	57%	71%	76%
	7%	4%	28%	33%	10%	13%
2. LO with SAC	8%	33%	15%	0%	8%	6%
(Prevalence: 11%)	88%	66%	4%	1%	54%	0%
	4%	1%	81%	99%	37%	94%
3. LO with SAC only in additive tasks	3%	25%	8%	2%	9%	2%
	92%	73%	2%	0%	79%	98%
(Prevalence: 38%)	5%	2%	90%	98%	12%	0%
4. Correct meaning tasks with LO/SAC elsewhere	89%	82%	23%	7%	11%	2%
	0%	13%	3%	0%	79%	85%
(Prevalence: 8%)	11%	5%	74%	93%	10%	13%
5. LO apart from additive tasks	16%	24%	91%	39%	29%	8%
(Prevalence: 15%)	75%	70%	9%	17%	64%	89%
	9%	6%	0%	44%	7%	3%
6. Mostly correct	61%	86%	69%	31%	85%	64%
(Prevalence: 5%)	30%	12%	10%	7%	14%	33%
	9%	2%	21%	62%	1%	3%

Figure 3: The six latent classes with their item-response probabilities for every task

Class 2 (prevalence: 11%) is characterised by high probabilities for SAC responses. However, for *Biros* a LO response is more likely than a SAC response. In the meaning tasks that do not offer a SAC option, LO is most likely; a correct response in *Bricks* is possible. Since this is the class with the highest probability for SAC in proportional tasks, students in this class seem to be quite convinced that coefficients stand for (possible) solutions also in different equation types. We label this class "*LO with SAC*".

Class 3 (prevalence: 38%) is characterised by very high probabilities for SAC responses in additive tasks and high probabilities for LO responses in all other tasks. This indicates a rather consistent interpretation of letters as abbreviations for involved objects in combination with an interpretation of coefficients as solutions in additive equations. We label this class *"LO with SAC only in additive tasks*".

Class 4 (prevalence: 8%) is characterised by high probabilities for correct responses in meaning tasks, high probabilities for SAC in additive tasks, and high probabilities for LO in proportional tasks. Students in this class seem to be able to identify the correct meaning of an algebraic letter when directly being asked for it. However, when choosing equations they still fall into the trap of interpreting letters as abbreviations (and coefficients as solutions in additive equations). We label this class "*Correct meaning tasks with LO/SAC elsewhere*".

Class 5 (prevalence: 15%) is characterised by medium to high probabilities for LO in all tasks other than additive tasks. While in *Garden* the correct response is most likely, in *Wheels* SAC is slightly more likely than the correct response. For this response pattern, it is impossible to identify one reason (see Discussion). We label this class "*LO apart from additive tasks*".

Class 6 (prevalence: 5%) is characterised by correct responses being most likely in almost all tasks. Only in *Wheels*, SAC is double as likely as the correct response. In *Ducks* and *Racetrack*, LO is also possible but half as likely as the correct response. This indicates at least a partial understanding that algebraic letters do not stand for abbreviations. We call this class "*Mostly correct*".

Discussion and Outlook

Utilising LCA, six distinct response pattern classes were identified, offering detailed insights into the relationship between students' comprehension, misconceptions, and test tasks. These classes play a crucial role in enhancing our understanding of how students interpret algebraic letters across various contexts. It is important to note that a comprehensive understanding of the implications is an ongoing research process, and this discussion marks our initial attempt at exploring these insights.

Starting with classes that exhibit at least some correct answers, a notable discovery is that Class 6, characterised by mostly correct answers, has a low prevalence of 5% and still shows many SAC responses to the *Wheels* task. The absence of a class labelled 'All answers correct' is not surprising, as only 21 students (1%) would belong to this class, which contradicts the principle of a good LCA model that avoids very rare classes. Class 4 (8%) is intriguing, displaying high probabilities for correct meaning tasks, but struggles when translating this understanding into equations. These students seem to possess a superficial knowledge of variable meanings, adequate for direct inquiries about meaning with one variable but insufficient when dealing with equations involving two variables. This underscores the importance of recognising that merely asking about the meaning of letters in simple contexts does not necessarily imply a deep and accurate understanding. Class 5 (15%) is characterised by a very high probability of a correct answer on the *Garden* task and LO/SAC in most other tasks. Since these students do not seem to grasp the meaning of letters, it is likely that these correct responses are not a result of (partial) understanding but of a strategy of combining given letters and numbers according to the described situation without proper understanding.

Regarding the LO misconception and its subtypes, it is crucial to highlight that the LU misconception had a minimal occurrence in the meaning tasks. Some students consistently show LO (Class 1, 23%); however, even in this class, the subtype SAC has a probability of 33% in *Wheels*. This might indicate that this task especially fosters students' urge to come up with a numerical solution. In general, the subtype SAC is often clearly present in additive tasks only (especially Classes 3 and 4). Such additive equations can probably be read more intuitively as a solution sentence such as "35 bikes plus 10 trikes have 100 wheels altogether" compared to proportional tasks that would have to be read as something like "Sam bought 30 biros in 10 packs". It is also possible that the additive tasks more easily trigger some students' desire to give a solution than proportional tasks. In this respect, Class 2 is exceptional: the probability for SAC is very high in *Racetrack,* but only 37% in *Biros*, while they are both proportional tasks (with similar response rates, see Figure 2). This might indicate that a SAC interpretation in proportional equations is more likely when the letters involved refer to non-physical objects (e.g., rounds in *Racetrack)* or can be interpreted as units (e.g., minutes in *Racetrack).*

The analysis of the identified classes underlines how important the task type, complexity, and context – including the realness of involved entities and the underlying structure of the equation – seems to be for correctly interpreting algebraic letters. These are aspects that need to be taken into account for teaching as well as assessment. For example, problems that focus on a numerical interpretation (like "Think of a number") or require operations on both sides of the equation might help to support a correct interpretation of the equal sign as well as of algebraic letters. Two further questions remain for future research: How do students transition from these classes after a lesson series about algebraic letters? And: How can this analysis improve the SMART-test *Meaning of Letters* diagnosis?

Acknowledgements

This research was supported by a Travel Award from the Interdisciplinary Centre for Educational Research (IZfB) of the University of Duisburg-Essen granted to Katrin Klingbeil.

References

Akhtar, Z., & Steinle, V. (2017). The prevalence of the 'letter as object' misconception in junior secondary students. In A. Downton, S. Livy & J. Hall (Eds.), *Proceedings of the 40th annual conference of the Mathematics Education Research Group of Australasia* (pp. 77–84). MERGA.

Arcavi, A., Drijvers, P., & Stacey, K. (2017). *The learning and teaching of algebra: Ideas, insights, and activities*. Routledge.

Brandenburger, M., Schwichow, M. (2023). Utilizing Latent Class Analysis (LCA) to Analyze Response Patterns in Categorical Data. In X. Liu, & W. J. Boone (Eds.), *Advances in Applications of Rasch Measurement in Science Education. Contemporary Trends and Issues in Science Education*, vol 57. (pp. 123 –154). Springer, Cham. https://doi.org/10.1007/978-3-031-28776-3_6

Clement, J., Lockhead, J., & Monk, G. (1981). Translation difficulties in learning mathematics. *American Mathematical Monthly*, *88*(4), 286–290.

Klingbeil, K., Rösken, F., Barzel, B., Schacht, F., Stacey, K., Steinle, V. & Thurm, D. (2024). Validity of multiple-choice digital formative assessment for assessing students' (mis)conceptions: Evidence from a mixed-methods study in algebra. *ZDM – Mathematics Education*. https://doi.org/10.1007/s11858-024-01556-0

Küchemann, D. (1981). Algebra. In K. M. Hart, M. L. Brown, D. E. Küchemann, D. Kerslake, G. Ruddock, & M. McCartney (Eds.), *Children's understanding of mathematics: 11–16* (pp. 102–119). John Murray.

MacGregor, M., & Stacey, K. (1997) Students' understanding of algebraic notation: 11-16. *Educational Studies in Mathematics*, *33*(1), 1–19.

Steinle, V., Gvozdenko, E., Price, B., Stacey, K., & Pierce, R. (2009). Investigating students' numerical misconceptions in algebra. In R. Hunter, B. Bicknell & T. Burgess (Eds.), *Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 491–498). MERGA.