Title
Developmental changes in students' use of dimensional comparisons to form ability self-concepts in math and verbal domains

Journal
Child Development, 94(1)

ISSN
0009-3920

Authors
Wan, Sirui
Lauermann, Fani
Bailey, Drew H
et al.

Publication Date
2023

DOI
10.1111/cdev.13856

Peer reviewed

DOI: 10.1111/cdev.13856

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

Developmental changes in students' use of dimensional comparisons to form ability self-concepts in math and verbal domains

Sirui Wan[1,2] | Fani Lauermann[3] | Drew H. Bailey[2] | Jacquelynne S. Eccles[2]

[1]University of Wisconsin-Madison, Madison, Wisconsin, USA
[2]University of California, Irvine, Irvine, California, USA
[3]TU Dortmund University, Dortmund, Germany

Correspondence
Sirui Wan, Department of Psychology, University of Wisconsin-Madison, 1202 W. Johnson St., Madison, WI 53706, USA.
Email: swan25@wisc.edu

Funding information
American Psychological Association; Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: HD17553; Jacobs Foundation; National Science Foundation, Grant/Award Number: DRL-1108778 and HRD-1231347

Abstract

Dimensional comparisons (i.e., comparing own performances across domains) may drive an increasing differentiation in students' math and verbal self-concepts over time, but little longitudinal research has directly tested this assumption. Using cross-sequential data spanning Grades 1–12 (N = 1069, ages 6–18, 92% White, 2% Black, 51% female, collected 1987–1996), this study charted age-related changes in the role of dimensional comparisons in students' ability self-concept formation. It used three types of self-concept measures: peer comparisons, cross-domain comparisons, and no comparisons. Results indicated that the increase in students' use of dimensional comparisons in self-evaluations substantially contributed to the increasing differentiation in students' math and verbal self-concepts over time. Findings highlight the importance of dimensional comparisons in the development of students' ability self-concepts.

Students' ability self-concepts—that is, their self-perceptions of what they are good at and how good they are at different types of tasks—predict their future domain-specific achievements and choices (Eccles & Wigfield, 2020). For instance, compared to students who doubt their math ability or aptitude, those who perceive themselves as good at math are more likely to engage in math class (Cai et al., 2018), earn good grades in math (Marsh et al., 2005), enroll in advanced math classes as they progress throughout their educational careers (Simpkins et al., 2006), and pursue postsecondary and occupational goals in math-intensive fields (Lauermann et al., 2017). Ability self-concepts are domain-specific, such that individuals can feel good about their reading ability but have a low self-concept of math ability. Although students' math and verbal achievements are consistently positively correlated over time (Möller et al., 2020), their math and verbal ability self-concepts become increasingly differentiated over the school years, which is reflected in a decreasing correlation between these constructs over the school years (Wan et al., 2021). This raises the question of what information students use to form their ability self-concepts in different domains and across different stages of their educational careers, and whether students' use of different sources of information to determine their relative strengths and weaknesses across domains might contribute to their increasingly differentiated self-concepts of ability over time.

One way students form their domain-specific ability self-concepts is through dimensional comparisons, that is, students compare their levels of performance in different domains such as math and language arts to determine their relative strengths or weaknesses in these domains (Möller & Marsh, 2013). A student may think, for instance, "Compared to my ability in language arts, I am not very good at math." As a consequence of these dimensional comparisons, high levels of performance

Abbreviations: I/E, internal-external frame of reference; RQ, research question.

in one domain (e.g., language arts) negatively predict students' ability self-perceptions in contrasting domains (e.g., math), even when same-domain achievement differences (e.g., in math) are controlled for. Additional sources of information that shape students' self-concepts of ability are social comparisons (i.e., comparing one's own performance in a given domain with that of relevant others in the same domain) and temporal comparisons (i.e., comparing one's current performance with past performances in the same domain). Although all three types of comparisons are fundamental for students' ability self-concept formation, only dimensional comparisons are presumed to be a key contributor to the increasing differentiation of students' math and verbal self-concepts over the school years (Wan et al., 2021). Indeed, meta-analytic evidence suggests that, after controlling for same-domain achievement differences, the predictive effects of students' math (vs. reading) achievement on their reading (vs. math) self-concept are more negative for older than younger students (Möller et al., 2020). Thus, older students appear more likely to use dimensional comparisons than younger ones. However, these analyses present only indirect evidence of dimensional comparisons, which are inferred from the associations between students' cross-domain achievements and ability self-concepts. In this literature, students' use of dimensional comparisons has rarely been directly assessed, for instance, by asking students to respond to comparative statements about their abilities across domains at different stages of their educational careers. As a result, there is also no direct longitudinal evidence that students increasingly use dimensional comparisons, corresponding to the increasing differentiation of students' math and verbal self-concepts over the school years.

In the present study, we use longitudinal data from the Childhood and Beyond study, which includes items that instructed students to evaluate their math and verbal ability self-concepts over time either (a) without explicit instructions about which type of comparison to use (i.e., *non-referenced* ability self-concepts such as: "How good are you at math [reading]?"), (b) based on dimensional comparisons (i.e., *internally referenced* ability self-concepts such as: "How good are you at math [reading], compared to other activities and subjects?"), or (c) based on social comparisons (i.e., *externally referenced* ability self-concepts such as: "If you were to list all the students from best to worst in math [reading], where are you?"). By comparing students' answers across these items, we provide direct evidence for the increase in students' use of dimensional comparisons in their ability self-evaluations. Furthermore, by examining the links between students' internally referenced self-concepts in math and verbal domains longitudinally, we examine whether an increase in students' use of dimensional comparisons is associated with a corresponding increase in cross-domain differentiation of students' self-concepts over the school years.

We focus on age-related changes in students' use of dimensional comparisons and whether these changes can explain the increasing differentiation of students' math and verbal self-concepts over time. Studying these questions is important because it clarifies the role of dimensional comparisons in shaping students' self-evaluations at different stages in their educational careers, and because dimensional comparisons can significantly affect students' future educational choices and areas of specialization. Our study is the first to chart developmental changes based on direct assessments of students' use of dimensional comparisons from Grades 1 through 12.

Dimensional comparisons in ability self-concept formation

Dimensional comparisons (Möller & Marsh, 2013), along with social (Festinger, 1954) and temporal comparisons (Albert, 1977), are proposed as a key influencing factor in the formation of students' ability self-concepts (Skaalvik & Skaalvik, 2002). According to the internal-external frame of reference (I/E) model (Marsh, 1986) and dimensional comparison theory (Möller & Marsh, 2013), dimensional comparisons involve comparing one's performance in a given school subject with own performances in other subjects. Dimensional comparisons between academic domains that are perceived as dissimilar, such as math and language arts, are assumed to lead to negative contrast effects in the formation of students' ability self-concepts in these domains. For instance, relatively high levels of achievement in one domain (e.g., language arts) set a high standard against which students' compare their ability in the contrasting domain (e.g., math). Due to dimensional comparisons, students who perform better in reading than in math tend to have lower self-perceptions of math ability than students with identical math ability but lower reading performance. A large body of research has provided support for the existence of dimensional comparison effects in students' ability self-concept formation (Möller et al., 2020).

Students' use of dimensional comparisons has been proposed as a key reason for the discrepancy in the correlations between students' math and verbal achievements versus corresponding math and verbal self-concepts of ability (Marsh, 1986). Specifically, whereas students' abilities in math and verbal domains are positively correlated, their math and verbal self-concepts are only weakly or not at all related (Möller et al., 2009, 2020). In other words, due to dimensional comparisons, students' math and verbal self-concepts become increasingly differentiated despite the strong positive correlation between their math and verbal achievements.

Age-related changes in dimensional comparisons

The differentiation between students' math and verbal self-concepts increases over the school years (Möller et al., 2020; Wan et al., 2021). In a meta-analysis, Wan et al. (2021) found that the correlation between students' math and verbal self-concepts flipped from a positive value to a negative one during the adolescent years. They further suggested that students' increasing reliance on dimensional comparisons over age is likely the main reason behind this "correlation flip." Indeed, many studies have reported negative (e.g., Gunderson et al., 2017; Lauermann et al., 2020; Umarji et al., 2018) and close-to-zero correlations (e.g., Möller et al., 2011; Wolff et al., 2020) between students' math and verbal self-concepts among secondary school students. In contrast, the evidence for elementary school students—although comparatively less prolific and less consistent—points to positive associations in the early school years (e.g., Schneider & Sparfeldt, 2020; Weidinger et al., 2019). Furthermore, dimensional comparison effects become stronger not only in secondary school but also across the elementary school years (Ehm et al., 2014; Lohbeck & Möller, 2017; Schneider & Sparfeldt, 2020; Sewasew & Schroeders, 2019; Weidinger et al., 2019).

Theoretically, it is sensible that students' use of dimensional comparisons increases over age due to important changes in students' reasoning about competence and the nature of their educational environments (Cimpian, 2017; Dweck, 2002; Eccles et al., 1993; Harter, 2012; Higgins & Eccles-Parsons, 1983; Nicholls, 1978; Parsons & Ruble, 1977; Stipek & Mac Iver, 1989). First, young students may not be able to make dimensional comparisons, both due to their stage of cognitive development and the external sources of information that are available to them. In particular, dimensional comparisons require the cognitive ability to differentiate different levels of competence across domains, as well as the ability to compare and integrate different sources of information across domains and over time (Stipek & Mac Iver, 1989; Weidinger et al., 2019). Moreover, dimensional comparisons are also affected by the learning environment and by normative evaluations across subjects (e.g., the increasing emphasis on normative performance feedback and grades over the school years, Eccles & Roeser, 2011). Second, students may become increasingly motivated to make dimensional comparisons because of the need (or pressure) to choose among and specialize in different domains in later school years (Gaspard et al., 2020). Self-differentiation, that is, the desire to identify one's relative strengths and weaknesses across different domains, has emerged as a key motivation behind making dimensional comparisons in high school samples (Wolff, Helm, & Möller, 2018). Dimensional comparisons support students' decision-making about which course to enroll in or which major and career paths to choose when such decisions need to be made in secondary education.

Direct and indirect assessments of dimensional comparisons

To date, researchers have mainly relied on a path-analytic approach to study dimensional comparison effects in different populations and age groups (for a meta-analysis, see Möller et al., 2020). As noted previously, numerous studies show that students' achievement in one domain (e.g., language arts) negatively predicts students' ability self-concept in the contrasting domain (e.g., math), when same-domain achievement differences (e.g., in math) are statistically controlled. The estimates of these negative cross-domain paths from achievement in one domain to ability self-concept in a contrasting domain are interpreted as evidence of dimensional comparison effects.

A limitation of the path-analytic approach and its interpretation is that whether individuals engage in dimensional comparisons is not directly measured. In most studies, students' academic self-concepts are assessed based on non-referenced items (e.g., "How good are you at math?"). These items do not ask students to use a specific internal or external frame of reference for ability self-evaluations. Studies using such non-referenced items often show the existence of both social and dimensional comparison effects, suggesting that students are likely to spontaneously use some frame of reference (e.g., peers, other subjects) to answer these questions even when they are not explicitly prompted to do so (Marsh et al., 2019). Using this approach, we can only make indirect inferences about the effects and the type of comparison students engage in when they evaluate their ability self-concepts. There are a few experimental studies directly assessing the effects of dimensional comparisons on individuals' self-evaluations (Möller & Köller, 2001; Müller-Kalthoff, Jansen, et al., 2017; Pohlmann & Möller, 2009; Strickhouser & Zell, 2015). However, all of them focused on specific age groups (either high school or college students), and are thus not suited for answering developmental questions, which are of primary interest for the present study.

A complementary and more direct way of capturing dimensional and social comparison processes is to implement separate assessments that elicit students' comparative evaluations of their ability either in reference to alternative domains (i.e., dimensional comparisons) or in reference to relevant others (i.e., social comparisons). Several cross-sectional studies have applied such an approach (Bong, 1998; Dickhäuser, 2005; Marsh & Yeung, 2001; Müller-Kalthoff, Helm, et al., 2017; Skaalvik & Skaalvik, 2004; Wolff, 2020; Wolff, Helm, Zimmermann, et al., 2018). These studies show that explicitly asking students to use either internal or external frames of reference results in different ratings of students' ability self-concepts. The correlations between math and verbal self-concepts were negative when using dimensional comparisons (i.e., an internal frame of reference), but positive when using social comparisons (i.e., an external frame of reference). Accordingly, internally and externally referenced items do seem to capture different comparison processes in self-evaluations. However, due to their cross-sectional design, these studies are not suitable for analyses of developmental changes in students' use of dimensional versus social comparisons.

The present study

The current longitudinal study expands upon previous research by using direct assessments of dimensional and social comparisons across Grades 1–12 that explicitly asked students' to make ability self-evaluations based on internally referenced (i.e., comparisons across different subjects) or externally referenced items (i.e., comparisons with other students). The design of the study makes it possible to elicit students' use of dimensional comparisons more directly than inferring such comparisons from correlations between students' math and reading self-concepts, as assessed by non-referenced items.

The present study has two main objectives. First, we test whether students' use of dimensional comparisons in forming ability self-evaluations increases over the school years. To this end, we examine possible changes in the predictive effects of internally referenced ability self-evaluations (e.g., "How good are you at math, compared to other activities and subjects?") and externally referenced ability self-evaluations (e.g., "If you were to list all the students from best to worst in math, where are you?") on non-referenced self-concept items (e.g., "How good are you at math?"). These analyses allow us to determine whether the relative weight of students' use of dimensional relative to social comparisons in the formation and prediction of students' (non-referenced) domain-specific ability self-concepts increases over time.

Second, we examine whether an increase in students' use of dimensional comparisons corresponds to an increasing differentiation of students' math and verbal self-concepts over the school years. Following Wan et al. (2021), we examine whether the strength of the correlations between students' math and verbal ability self-evaluations changes from positive in the early school years to zero or negative in later years, which would indicate that these self-evaluations are initially well-aligned but become more differentiated over time. We expand upon prior evidence by examining potential changes in the correlations between students' (a) non-referenced, (b) externally referenced, or (c) internally referenced math and verbal ability self-evaluations longitudinally over the school years. If dimensional comparisons are driving an increasing cross-domain differentiation over time, then we would expect the strongest decline in these math-verbal associations for students' internally referenced ability self-evaluations.

The corresponding research questions (RQs) and expected results are:

RQ1: How does the relation between students' internally referenced and non-referenced self-concepts change over time in the domains of math and language arts? For both domains, we expected that, after controlling for the predictive effects of externally referenced self-concepts, this association should increase over time. Given evidence on children's social-cognitive development, the increasing salience and prevalence of normative evaluations across subjects, and the pressure to specialize in later school years, we predict that older students will become increasingly likely to use dimensional comparisons to form their ability self-concepts over the school years.

RQ2a: How does the relation between students' non-referenced math and verbal self-concepts change over time? We expect that the correlation between students' non-referenced self-concepts in math and language arts will decrease over time, flipping from a positive value to a zero or negative one during the adolescent years (i.e., showing an increasing differentiation, see Wan et al., 2021).

RQ2b: How does the relation between students' externally referenced math and verbal self-concepts change over time? Based on findings from previous studies (Dickhäuser, 2005; Marsh & Yeung, 2001; Wolff, Helm, Zimmermann, et al., 2018), we expect that the correlation will be moderately positive. We also predict that it will remain positive over the school years because students' math and language achievements are consistently positively correlated across development (Möller et al., 2020).

RQ2c: How does the relation between students' internally referenced math and verbal self-concepts change over time? We expect a negative correlation between students' internally referenced math and language self-concepts in secondary school, consistent with prior evidence (Dickhäuser, 2005; Marsh & Yeung, 2001; Wolff, Helm, Zimmermann, et al., 2018). This hypothesis is less certain for students in the early school years due to a lack of prior evidence using direct assessments of dimensional comparisons. We expect an age-related decrease in this correlation (i.e., the magnitude of the negative correlation should increase over age) due to students' increasing ability and tendency to differentiate their ability self-evaluations (i.e., identify their relative strengths and weaknesses across different domains).

Due to the richness of prior research on dimensional comparisons, we were able to formulate theory- and evidence-based expectations for all analyses. Thus, the present analyses are confirmatory in nature. They serve to address the specific hypotheses we have outlined above for each RQ.

METHOD

Participants and procedure

We used data from the Childhood and Beyond study, an existing longitudinal study of the development and socialization effects on children's achievement motivation and behavior. Data were collected from three cohorts of children and their parents and teachers between 1987 and 1996, beginning when Cohort 1 was in kindergarten, Cohort 2 was in Grade 1, and Cohort 3 was in Grade 3. These children were initially followed for three consecutive years. After a three-year gap in funding, additional information was collected from the students for three more consecutive years. For a detailed description, see http://garp.education.uci.edu/cab.html.

We used data from Waves 2, 3, 4, 5, and 7 in this study because those data points had students' reports of their ability self-concepts in both math and language domains over a 12-year period. Thus, data were collected in Grades 1, 2, 3, 7, and 9 for Cohort 1, Grades 2, 3, 4, 8, and 10 for Cohort 2, and Grades 4, 5, 6, 10, and 12 for Cohort 3 (see Table 1). We excluded data from students who had never reported on their ability self-concepts across the five waves of data collection. The final sample consisted of 1069 children, of whom 318 were in Cohort 1 (154 girls and 164 boys), 330 were in Cohort 2 (171 girls, 159 boys), and 421 were in Cohort 3 (220 girls and 201 boys).

TABLE 1 Sample description by cohort, grade level, and age at each time point of data collection

	Cohort 1	Cohort 2	Cohort 3
Time 1: 1987–1988			
Grade level	Grade 1	Grade 2	Grade 4
N	294	314	264
Mean age in years (*SD*)	6.76 (.38)	7.73 (.39)	9.70 (.36)
Time 2: 1988–1989			
Grade level	Grade 2	Grade 3	Grade 5
N	281	298	401
Mean age in years (*SD*)	7.76 (.39)	8.72 (.38)	10.70 (.38)
Time 3: 1989–1990			
Grade level	Grade 3	Grade 4	Grade 6
N	244	251	372
Mean age in years (*SD*)	8.76 (.38)	9.71 (.37)	11.70 (.36)
Time 4: 1993–1994			
Grade level	Grade 7	Grade 8	Grade 10
N	184	190	276
Mean age in years (*SD*)	12.75 (.38)	13.73 (.38)	15.69 (.35)
Time 5: 1995–1996			
Grade level	Grade 9	Grade 10	Grade 12
N	151	151	195
Mean age in years (*SD*)	14.73 (.37)	15.71 (.36)	17.67 (.34)

Note: The average age was computed as of January 1 of each year shown (e.g., January 1, 1996).

Children attended 10 public, elementary schools in four middle-class school districts in the suburbs of a large midwestern city in the United States. Among those who reported their ethnicity, the sample was primarily European American (91.9%), with a very small minority of African Americans (1.5%), Asians (5.7%), American Indians (<1%), and Hispanics (<1%). Two hundred and fifteen students did not report their ethnicity. Gender was almost perfectly balanced across all waves of data collection. In general, the families were middle- or working-class, two-parent families (90% two parents). Family income in 1986 ranged from \$10,000 to over \$80,000, with an average income between \$40,000 and \$50,000.

Measures

All survey items were developed by Eccles and colleagues, that is, for ability self-concepts or expectancies, interest value, and importance value (see Wigfield & Eccles, 2000). The math and language ability self-concept measures each included five items. For the math domain, the same items were asked in all waves of data collection. For the language domain, the items referred to reading in Waves 2–4 and to English in later waves to account for changes in the content and the main focus of the language arts classes after elementary school. In other words, the items were tailored to be relevant to what students were learning in school at each stage (e.g., elementary-level students had a subject called reading but no class called English, and vice versa for students in Grades 7–12). These self-concept items were divided into three categories based on different frames of reference: non-referenced (i.e., no frames of reference in the prompt of measurement so that students may freely choose frames of reference for comparisons), externally referenced (i.e., based on social comparisons with peers), and internally referenced (i.e., based on intra-individual dimensional comparisons across domains) items.

Across all waves, students answered questions about the math domain (e.g., math ability self-concept, math interest, math anxiety) first, and questions about the language arts domain later in the survey. The order in which questions about students' ability self-concepts were presented changed between Waves 2–4 (i.e., Times 1–3) and Waves 5 and 7 (i.e., Times 4–5). At all time points, these questions started with a non-referenced item ("How good at math/reading/English are you?") and ended with two additional non-referenced items. For Times 1–3, the externally referenced item was presented before the internally referenced item. For Times 4 and 5, the internally referenced item was presented before the externally referenced item.

Non-referenced math/language ability self-concept items

Three non-referenced items were used in the analyses: "How good at math [reading/English] are you?" (1 = not very good, 7 = very good). "How well do you expect to do in math [reading/English] this year?" (1 = not at all well, 7 = very well); "How good would you be at learning something new in math [reading/English]?" (1 = not very good, 7 = very good). The average scores of the three items (for each domain) were used in the analysis. The internal consistency of these items was satisfactory for most time points (math: $\alpha = [.49, .93]$, language: $\alpha = [.60, .93]$). The average test–retest reliability for adjacent waves of data collection for these items was $\bar{r} = .42$ for math and $\bar{r} = .39$ for language.

Externally referenced math/language ability self-concept items (i.e., based on social comparisons)

Students' externally referenced self-concepts were assessed with the item: "If you were to list all the students from best to worst in math [reading/English] where are you?" (1 = one of the worst, 7 = one of the best). The average test–retest reliability for adjacent waves of data collection for these single-item measures was $\bar{r} = .39$ for math and $\bar{r} = .40$ for language.

Internally referenced math/language ability self-concept items (i.e., based on dimensional comparisons)

Students' internally referenced self-concepts were assessed with the item: "Compared to most of your other school subjects, how good are you at math [reading/English]?" (1 = a lot worse, 7 = a lot better). The average test–retest reliability for adjacent waves of data collection for these single-item measures was $\bar{r} = .37$ for math and $\bar{r} = .34$ for language.

Teacher-evaluated math/reading aptitude in elementary school

Elementary school teachers of participating students evaluated their students' math and reading aptitude in the first four waves of data collection (kindergarten through Grade 6) using two items: "Compared to other children, how much innate ability or talent does this child have in math [reading]?" ranging from 1 (very little) to 7 (a lot), and "How well do you expect this child to do next year in math [reading]?" ranging from 1 (very poorly) to 7 (exceptionally well). The average scores of the two items (for each domain) were reported in Table S1. The internal consistency of these ratings was very good (math: $\alpha = [.82, .91]$, language: $\alpha = [.81, .90]$).

Analysis

Preparatory tests of the validity of externally- versus internally referenced single-item measures

Before answering our main RQs, we conducted path analyses and used teacher ratings of student aptitude to validate the externally and internally referenced assessments of students' ability self-concepts. This approach is based on the Extended I/E Model (Dickhäuser, 2005; Marsh & Yeung, 2001) and is shown in Figure 1 (Model A). As noted previously, negative cross-domain predictive effects of students' math (or verbal) achievement on their verbal (or math) ability self-concepts, and controlling for same-domain achievement differences, have been interpreted as a sign of students' use of dimensional comparisons. Accordingly, in a set of preparatory analyses, we tested such cross-domain predictive effects of teacher-rated student ability in math and reading on students' internally versus externally referenced self-concepts of ability in these two domains. For example, students with low performance in reading should be more likely than those with high performance in reading to think that *they are better in math than in other school subjects* (measured by *internally* referenced items). However, we do not have a priori predictions about whether students with low performance in reading will be more likely than those with high performance in reading to think that *they are better at math than are their classmates* (measured by

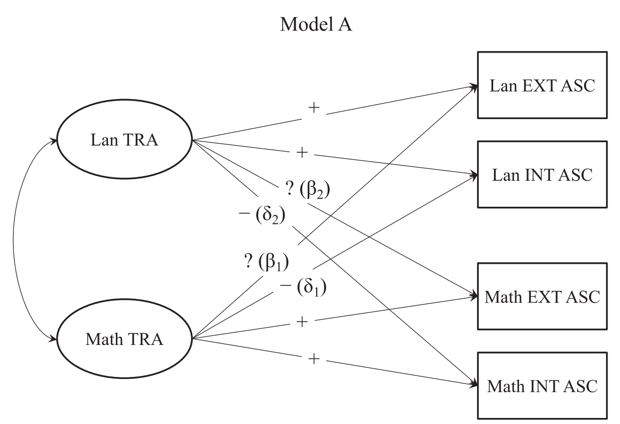

FIGURE 1 Conceptual model depicting the relations between achievement and frame-specific ability self-concepts according to Marsh and Yeung (2001). "+" and "−" refer to the predicted direction of the path coefficients, whereas "?" indicates that there are no a priori predictions. ASC, ability self-concept; EXT, externally referenced; INT, internally referenced; Lan, language arts; TRA, teacher-rated aptitude. β_1 (β_2) will be compared with δ_1 (δ_2) to validate that there are differences between the constructs measured by internally referenced versus externally referenced items.

externally referenced items). If the predictive effects of teacher-rated math (vs. reading) ability on students' *externally* referenced self-concepts in reading (vs. math) are different from the corresponding predictive effects of teacher-rated math (vs. reading) ability on students' *internally* referenced self-concept in reading (vs. math), then our internally referenced self-concept items are likely to capture a different process from the one captured by the externally referenced self-concept items. Stronger cross-domain effects should emerge for internally referenced than externally referenced items (see Dickhäuser, 2005). Wald tests were used to compare the strength of these predictive paths (see β_1 vs. δ_1 and β_2 vs. δ_2 in Figure 1).

Regression analysis for RQ1

To estimate the independent contributions from dimensional and social comparisons to students' ability self-concept formation, non-referenced ability self-concepts were regressed on both internally and externally referenced ability self-concepts for each cohort and at each wave. For each time point, internally and externally referenced ability self-concepts were entered in the regression models at the same time.

Structural equation modeling analysis for RQ2

To assess if the correlations between non-referenced math and language ability self-concepts change over time, we tested a quasi-simplex measurement model (Model B) presented in Figure 2. Similar models were used to examine the correlations between externally referenced math and language ability self-concepts (Model C) and the correlations between internally referenced math and language ability self-concepts (Model D). These analyses allowed us to examine the strength of associations between constructs of interest at different time points in students' educational careers. To adjust for the reliability of each single-item measure at each time point in the quasi-simplex models, the measurement errors were corrected by using the approach suggested by Anderson and Gerbing (1988). Specifically, we defined a single-item latent variable by fixing its factor loading to $.95 \times s$ (*SD* of the observed variable) and fixing its error variance to $.1 \times s^2$. Importantly, such error corrections rely on the assumption that error variance is truly "error" and are just a guess for how measurement error might have affected the results. We compared the results with and without such measurement error corrections to test the robustness of our findings. To empirically assess whether the developmental pattern is robust across cohorts in the dataset, we compared models in which the synchronous covariances between students' math and verbal ability self-concepts were constrained to be equal across the three cohorts if they are in the same grade (e.g., Grade 2 in Cohort 1 vs. 2; see Figure 2) to models in which all covariances were free to vary across the three cohorts. Model comparisons were based on changes in chi-square relative to changes in degrees of freedom (Kline, 1998).

All structural equation modeling analyses were run in Mplus Version 7.4 (Muthén & Muthén, 1998–2017). The full information maximum likelihood approach implemented in Mplus was used to deal with missing data. Maximum likelihood estimation with robust standard errors was used for parameter estimation to account for the violations of normality. The descriptive statistics and regression analyses were conducted in Stata/*SE* 14.0.

Robustness check

To rule out potential order effects for non-referenced items (e.g., students' potential tendency to answer the non-referenced items differently after having just filled out the internally or externally referenced items), we conducted additional analyses by using only the first of the three non-referenced items in all analyses. This item (i.e., "How good at math[reading/English] are you?") was always asked first, before internally or externally referenced items.

RESULTS

Descriptive statistics and zero-order correlations are shown in Tables S1–S4. These analyses show that teachers' ratings of students' ability in math versus reading are very highly correlated throughout elementary school (r = [.71, .82], see Figure S1), whereas the corresponding correlations between students' self-concepts decline during this same period (r = [−.29, .37]). These results are consistent with previous literature (Möller et al., 2009, 2020) and indicate that a gap emerges between students'n math-verbal ability versus math-verbal self-concept correlations during this time. This discrepancy in the correlational patterns between students' math-verbal achievements versus self-concepts is thus likely driven by dimensional comparisons rather than by students' increasingly differentiated abilities over this period. We also found that students reported higher means for their math and reading ability self-concepts in Grades 1–4 than in later school years (see Table S1).

Preliminary analyses: Validating internally and externally referenced assessments of students' ability self-concepts

To examine the extent to which frame-specific ability self-concept items capture students' dimensional

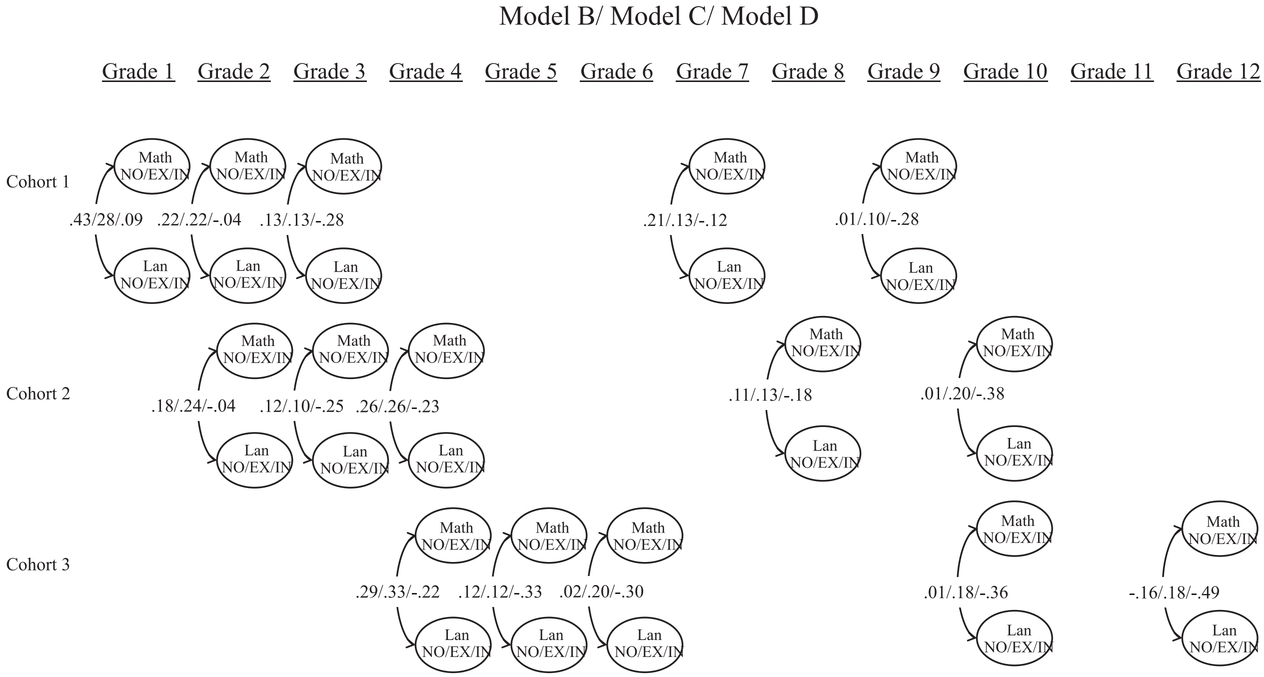

FIGURE 2 Estimates for quasi-simplex measurement model assessing the associations between students' math and language ability self-concepts over time, by cohort. Model B is for non-referenced items; Model C is for externally referenced items; Model D is for internally referenced items. EX, externally referenced; IN, internally referenced; Lan, language arts; NO, non-referenced. The models are saturated but only correlations that correspond to our research questions are shown for the sake of readability. The estimates in the figure are from the models that constrain the covariances between variables of interest for the same grade across different cohorts to be the same (i.e., Models B2, C2 and D2 in Table S5). Average scores of the three non-referenced items (for each domain) are used for Model B. Residuals are fixed at zero for latent variables in all the models.

comparison processes, we started with validating these items by linking them with teacher-evaluated math and reading aptitudes reported at the same time points. The student aptitude data are available for the elementary school years in Cohort 1 (Grades 1–3), Cohort 2 (Grades 2–4), and Cohort 3 (Grades 4–6). We tested the Extended I/E model (see Model A in Figure 1) in the nine available Cohort × Grade data sets, which we refer to as Models A1-A9. The Extended I/E model proposes that if the internally referenced self-concept items do indeed capture students' self-concept based on dimensional comparisons, then there should be negative paths from achievement (e.g., in math) to students' internally referenced self-concepts in the non-corresponding domain (e.g., language). In support of this hypothesis, students' math aptitude negatively predicted internally referenced self-concept items in reading, after controlling for students' reading aptitude (δ_1 = [−.52, −.10], see Table 2 and Figure S1). Similarly, students' reading aptitude negatively predicted internally referenced self-concept items in math, after controlling for students' math aptitude (δ_2 = [−.46, −.24], see Table 2; Figure S1). These results support the existence of dimensional comparison effects when internally referenced self-concept items were used.

Moreover, after controlling for the predictive effects of students' reading aptitude, the predictive effects of students' math aptitude on their internally referenced reading self-concept in Model A had larger effect sizes (see δ_1 in Table 2, average $\delta_1 = -.34$) than the corresponding predictive effects of math aptitude on students' externally referenced reading self-concept (see β_1 in Table 2, average $\beta_1 = -.17$). We found a similar pattern for the predictive effects of students' teacher-rated reading aptitude on their internally versus externally referenced math self-concept ratings (internal comparisons: average $\delta_2 = -.38$; external comparisons: average $\beta_2 = -.18$). Although the differences between δ and β do not reach statistical significance for some of the included time points (see *p*-values in Table 2), the pattern of δ having a larger effect size than β is consistent across all tested Models A1-A9. This pattern indicates that there are differences between the constructs measured by internally referenced versus externally referenced items. As expected, internally referenced items capture the dimensional comparison process to a greater extent.

RQ1: How does the relation between students' internally referenced and non-referenced self-concepts change over time?

Because we are interested in the extent to which students weigh dimensional comparisons in ability self-evaluations, non-referenced ability self-concepts were

TABLE 2 Parameter constraints for cross-domain effects from teacher-rated aptitudes to internal/external ability self-concepts in Models A1–A9

Cross-domain effects	δ	β	Wald test	p
δ_1 (math TRA → lan internal ASC) = β_1 (math TRA → lan external ASC)				
Model A1: Cohort 1 Grade 1	−.10 (.12)	.10 (.12)	.28 (.23)	.24
Model A2: Cohort 1 Grade 2	−.33 (.11)	−.23 (.10)	.22 (.20)	.25
Model A3: Cohort 1 Grade 3	−.42 (.13)	−.29 (.12)	.25 (.19)	.19
Model A4: Cohort 2 Grade 2	−.38 (.09)	−.05 (.10)	.54 (.14)	.000
Model A5: Cohort 2 Grade 3	−.39 (.11)	−.21 (.10)	.34 (.16)	.03
Model A6: Cohort 2 Grade 4	−.35 (.12)	−.25 (.12)	.14 (.14)	.32
Model A7: Cohort 3 Grade 4	−.18 (.14)	−.12 (.15)	.11 (.18)	.53
Model A8: Cohort 3 Grade 5	−.41 (.10)	−.25 (.09)	.27 (.11)	.01
Model A9: Cohort 3 Grade 6	−.52 (.13)	−.24 (.12)	.42 (.16)	.008
Average	−.34	−.17		
δ_2 (lan TRA → math internal ASC) = β_2 (lan TRA → math external ASC)				
Model A1: Cohort 1 Grade 1	−.24 (.12)	−.12 (.12)	.23 (.15)	.13
Model A2: Cohort 1 Grade 2	−.31 (.11)	−.12 (.11)	.34 (.18)	.06
Model A3: Cohort 1 Grade 3	−.46 (.12)	−.29 (.12)	.34 (.19)	.06
Model A4: Cohort 2 Grade 2	−.30 (.08)	−.26 (.08)	.17 (.13)	.20
Model A5: Cohort 2 Grade 3	−.45 (.11)	−.21 (.10)	.51 (.17)	.003
Model A6: Cohort 2 Grade 4	−.45 (.13)	−.24 (.13)	.41 (.20)	.04
Model A7: Cohort 3 Grade 4	−.35 (.11)	−.17 (.10)	.33 (.19)	.08
Model A8: Cohort 3 Grade 5	−.40 (.09)	−.09 (.09)	.51 (.13)	.000
Model A9: Cohort 3 Grade 6	−.42 (.10)	−.14 (.10)	.50 (.15)	.001
Average	−.38	−.18		

Note: Standard errors are in parentheses. Models A1–A9 are the same model (Model A in Figure 1) except that they are using different data (e.g., Cohort 1 Grade 1 is used in Model A1).

Abbreviations: ASC, ability self-concept; External, externally referenced; internal, internally referenced; Lan, language arts; TRA, teacher-rated aptitude.

regressed on internally referenced ability self-concepts for each academic domain. We controlled for the role of students' use of social comparisons by adding externally referenced ability self-concepts in these regression models. Thus, we were able to compare changes in the relative predictive effects of dimensional versus social comparisons. For the math domain, the amount of variance uniquely explained by the internally referenced items kept increasing during the school years (see Table 3; Figure 3). A similar pattern was replicated for the language arts domain, even though there were cohort-specific fluctuations in the regression estimates (see Table 4; Figure 3).

A different pattern was observed for the externally referenced items. Specifically, the amount of variance uniquely explained by the externally referenced items stayed relatively stable during the secondary school years (beyond Grade 6). We also found that for non-referenced ability self-concepts, the amount of variance explained by the externally referenced items was larger than that by the internally referenced items across all time points.

These results were robust when we used only the first non-referenced item instead of the averaged score of all three non-referenced items in all the analyses (see Figure S2). This item was always asked first and was thus not affected by internally or externally referenced items. Therefore, our findings seem to be robust to item order effects. As shown in Figure S2, the difference in the amount of variance explained by the externally versus internally referenced items became smaller over the school years, as dimensional comparisons gained in importance. These findings are consistent with the assumption that, as students grow older, they assign greater weight to dimensional comparisons in their (non-referenced) ability self-evaluations. This pattern is striking and points to differentiated developmental processes underlying students' use of dimentional versus social comparisons.

RQ2: How do the relations between students' math and verbal self-concepts change over time for non-, externally, or internally referenced items?

The results for RQ2 are presented in Figures 2 and 4, and in Table S6 . For the synchronous correlations between students' non-referenced math and language self-concepts of ability, we observed a "correlation flip" from .43 in

TABLE 3 Regression estimates of frame-specific self-concept ratings in math predicting non-referenced self-concept ratings in math

	Grade level											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Internal self-concept Cohort 1	.13**	.09**	.15***				.23***		.28***			
External self-concept Cohort 1	.34***	.38***	.45***				.53***		.53***			
Internal self-concept Cohort 2		.11***	.14***	.19***				.22***		.27***		
External self-concept Cohort 2		.46***	.35***	.51***				.46***		.55***		
Internal self-concept Cohort 3				.15***	.20***	.21***				.25***		.30***
External self-concept Cohort 3				.51***	.45***	.48***				.55***		.55***

Note: Blank = no data for analysis; internal = internally referenced; external = externally referenced. For each time point, both internally and externally referenced self-concept ratings were entered in one regression model.
**p<.01.; ***p<.001.

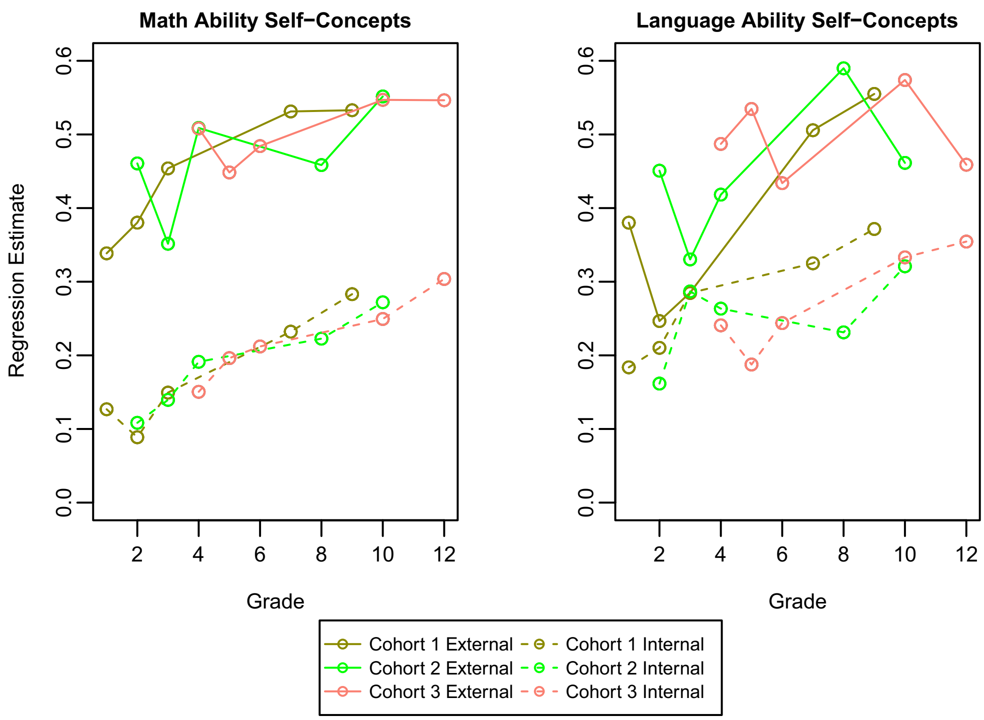

FIGURE 3 Regression estimates of frame-specific self-concept ratings predicting non-referenced self-concept ratings in math (left panel) and language (right panel). External, externally referenced; internal, internally referenced. For each time point, both internal and external self-concepts were entered in one regression model.

Grade 1 to −.16 in Grade 12 (see Model B in Figure 2), suggesting an increasing differentiation of students' math and verbal self-concepts. This pattern is robust if we use the first non-referenced item instead of the averaged score of all three non-referenced items in the analysis, ruling out potential item order effects (see Figure S3).

TABLE 4 Regression estimates of frame-specific self-concept ratings in language predicting non-referenced self-concept ratings in language

	Grade level											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Internal self-concept Cohort 1	.18***	.21***	.28***				.32***		.37***			
External self-concept Cohort 1	.38***	.25***	.29***				.51***		.56***			
Internal self-concept Cohort 2		.16***	.29***	.26***				.23***		.32***		
External self-concept Cohort 2		.45***	.33***	.42***				.59***		.46***		
Internal self-concept Cohort 3				.24***	.19***	.24***				.33***		.35***
External self-concept Cohort 3				.49***	.53***	.43***				.57***		.46***

Note: Blank = no data for analysis; internal = internally referenced; external = externally referenced. For each time point, both internally and externally referenced self-concept ratings were entered in one regression model.

$^{***}p < .001$.

Regarding the synchronous correlations between externally referenced math and language items (Model C in Figure 2), we found small and consistently positive correlations across all grade levels and in all cohorts. In contrast, the correlations between internally referenced math and verbal items shifted from positive to negative over time and across cohorts (Model D in Figure 2). For first- and second-graders, the estimated synchronous associations between internally referenced math and language ability self-concepts were positive or near-zero (standardized: ρ = [−.04, .09]). In Grade 3, the correlations between students' internally referenced math and language ability self-concepts became negative (standardized: ρ = [−.28, −.25]). As shown in Figure 4 and Table S6, there is a decline in the correlations between internally referenced items (i.e., an increase in the magnitude of the negative correlation) over the school years. These patterns are robust across cohorts as synchronous associations did not differ across cohorts for children in the same grades (see Table S5).

The different patterns in the results of the correlations between Model C and Model D indicate that externally and internally referenced items are capturing distinct developmental processes. They may also indicate that the decline in the correlations in Model D is not due to a measurement artifact that could have resulted, for instance, from young students having difficulty interpreting the items. In other words, if children have difficulty distinguishing between externally and internally referenced items, the patterns of results for Models C and D should be similar, which is not the case.

To interpret the age-related changes in the correlations between students' internally referenced math and language self-concepts, we illustrated all response patterns in the data in Figure 5. Younger students were far more likely than older students (e.g., beyond Grade 3) to rate themselves high in both math and language when being asked to use dimensional comparisons. As they grew older, students became more likely to rate themselves high in one domain but low in the other, suggesting that an increasing number of students perceived that they had relative strengths and weaknesses across domains.

DISCUSSION

Our study is the first in the literature to directly chart the changes in students' use of dimensional comparisons in self-evaluations of math and verbal abilities from Grades 1 to 12. Moreover, we provide evidence that the changes in students' use of dimensional comparisons largely drive the increasing differentiation of math and verbal self-concepts over time. Specifically, we found that the variances of ability self-concepts explained by dimensional comparisons increase over time, after controlling for social comparisons. In addition, we observed a decline in the correlations between math and language ability self-concepts over time, when the items explicitly asked students to evaluate their abilities based on dimensional comparisons. Our results are consistent with previous meta-analyses,

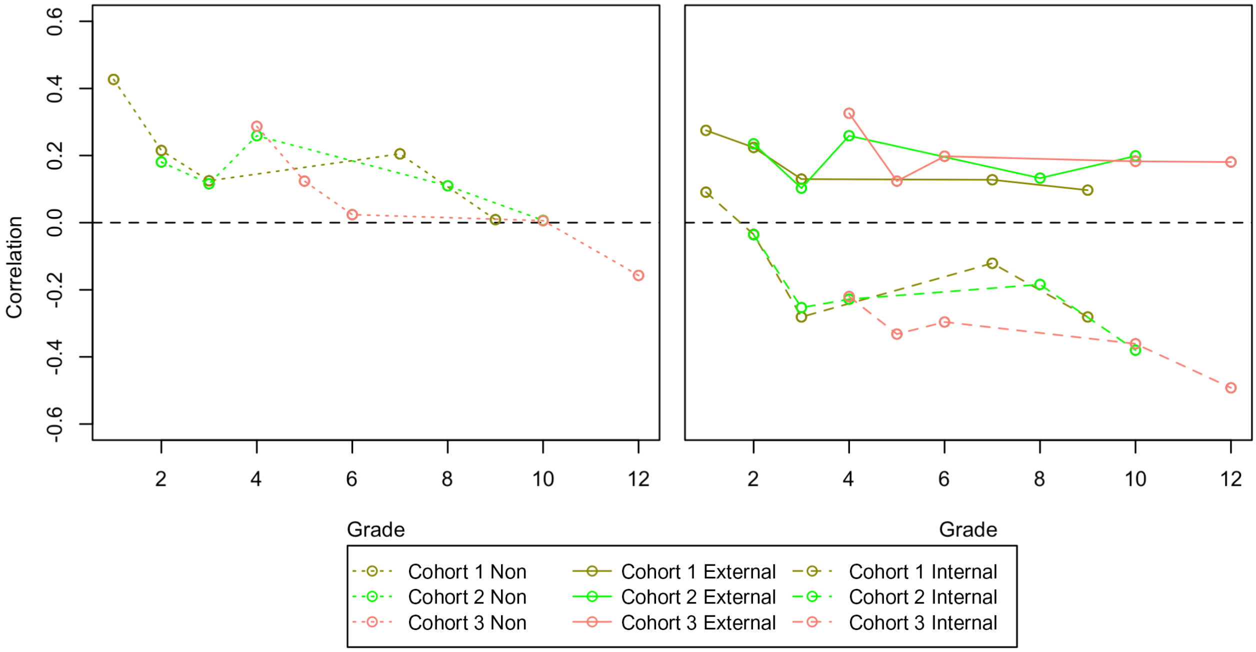

FIGURE 4 Correlations between math and language ability self-concepts. External, externally referenced; internal, internally referenced; non, non-referenced. The estimates in the figure are from the structural equation models that constrain the covariances between variables of interest for the same grade across different cohorts to be the same (i.e., Models B2, C2 and D2 in Table S5).

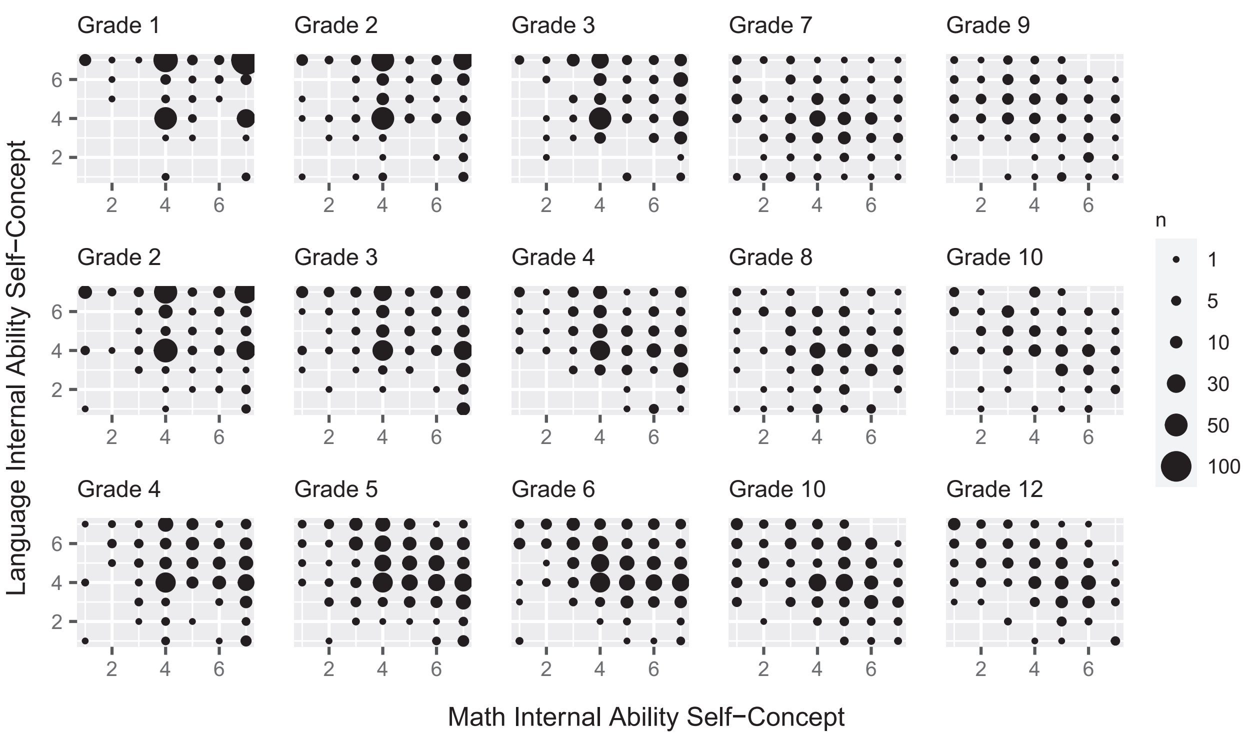

FIGURE 5 Weighted scatterplots of the correlations between math and language internal ability self-concepts (i.e., self-evaluations based on dimensional comparisons). The top panel is for Cohort 1, the middle panel is for Cohort 2, and the bottom panel is for Cohort 3. The size of a solid circle reflects the number of participants. Internal, internally referenced.

which suggested that there is an upward trend in dimensional comparison effects over time (Möller et al., 2020) and a flip in the correlation between students' math and verbal self-concepts during the adolescent years (Wan et al., 2021).

Changes in students' use of dimensional comparisons in ability self-evaluations

We find evidence for increases in children's use of dimensional comparisons to form math and verbal ability self-concepts. Given prior evidence that dimensional comparison effects are stronger among secondary schoolers than among elementary schoolers (Möller et al., 2020; Wan et al., 2021), this finding is consistent with our expectations. There are several possible reasons for the increases. First, based on their cognitive development and the low availability of external, normative information about student ability, students in the early school years may not be ready to make dimensional comparisons. Harter (2012) concluded that it is until the ages of 8–10 years old (Grades 2–4) that children start to realize that one's self-attributes can be both positive and negative. Younger children often showed all-or-none thinking, for instance, believing that one is "all good" at different age-appropriate skills. Consistent with this observation, our results showed that the students reported more positive self-evaluations (i.e., higher means) for their math and reading ability self-concepts in Grades 1–4 than in later school years (see Table S1). Mac Iver (1987) found that the frequency of grades provided in the math classroom is associated with students' heavy reliance on dimensional comparisons during math ability self-assessments in Grades 5 and 6. Thus, the increases in normative evaluations over the school years may both facilitate and prompt dimensional comparisons.

Second, older students may be more motivated to use dimensional comparisons. One type of motivation is self-differentiation motivation, that is, students' motivation to identify their strengths and weaknesses across various domains (Wolff, Helm, & Möller, 2018). Students are provided with more opportunities to make choices and are expected to specialize in certain domains in later school years. Thus, older students might be more motivated to engage in dimensional comparisons and rely on dimensional comparative information to form their ability self-concepts and support their decision-making for important choices such as which course to enroll in or which college major and career path to pursue. Another type of motivation for dimensional comparisons is self-protection. As normative evaluations and social comparison information increase over time, so does the negative impact of social comparisons on students' self-esteem (Harter, 2012). Accordingly, students' may use dimensional comparisons as a means to protect and enhance their self-images by shifting their attention to domains in which they perceive comparative strengths. Indeed, Möller and Husemann (2006) found that students use dimensional comparisons as a means to improve their mood.

Dimensional comparisons and increasing differentiation of students' math and verbal ability self-concepts

Recent meta-analytic studies suggest that students' math and verbal self-concepts become increasingly differentiated as students grow older (Möller et al., 2020; Wan et al., 2021). Consistent with this finding, our results showed that the correlations between non-referenced math and verbal self-concepts decreased over time. Furthermore, we identified different patterns of age-related changes in the correlations between students' math and verbal self-concepts when students were explicitly prompted to use either dimensional or social comparisons, or when no explicit internal or external reference point was provided. While the correlations between externally referenced math and language ability self-evaluations were small and consistently positive over the school years, the corresponding correlations between internally referenced math and language self-evaluations were mostly negative and the magnitude of the negative associations increased in the later school years. Previous cross-sectional studies (Dickhäuser, 2005; Marsh & Yeung, 2001; Wolff, Helm, Zimmermann, et al., 2018) reported positive correlations between externally referenced math and language ability self-evaluations and negative correlations between internally referenced self-evaluations. Our study expands upon this prior evidence by presenting a longitudinal analysis of developmental changes in these associations over a 12-year period.

Comparing the different patterns of age-related change in the correlations between students' math and verbal self-concepts (Figure 4) suggests that the increase in students' use of dimensional comparisons plays a major role in the increasing differentiation in students' math and verbal self-concepts across Grades 1–12. Because the correlations between students' math and language achievements are strong and consistently positive over time (Möller et al., 2009, 2020), alternative explanations such as changes in the subject content similarity between math and language arts are insufficient. Furthermore, we show different developmental patterns for students' use of dimensional versus social comparisons. Thus, our study provides evidence that there is an increasing differentiation of math and language self-concepts at the group level because, as they grow older, more and more students start to perceive themselves as being better at one particular domain (here: math or verbal) than at others.

Limitations and future directions

A potential shortcoming of the present study is that the measurements used to operationalize dimensional and social comparison processes were each assessed with a single item, which can limit these measures' content validity and reliability. However, single items often do provide valid and reliable assessments of psychological phenomena such as academic self-concepts and values (Beymer et al., 2021; Gogol et al., 2014; for an overview, see Allen et al., 2022). In this study, we address concerns regarding the reliability and validity of our single-item assessments in three ways. First, we validated these items by relating them to teachers' ratings of students' math and reading aptitude. Our results suggest good criterion validity for these single items (i.e., the dimensional comparison effects are larger for internally referenced items than for externally referenced ones; see the Preliminary Analyses section in Results). Second, these analyses replicate prior research on the effects of dimensional comparisons, which used multi-item scales and relied on a path-analytic approach (Möller et al., 2020). Third, we implemented measurement error corrections for single-item scales, as proposed by Anderson and Gerbing (1988). Our findings were consistent after implementing these measurement error corrections.

Another limitation is that internally referenced items asked students to use dimensional comparisons by comparing their math or verbal ability with "other subjects." Accordingly, when students answer these questions, they may think of subjects other than math and language arts, for instance, such subjects as sports and science. If this is indeed the case, our analyses might be underestimating the magnitude of the negative correlations between internally referenced items for older students because there are more school subjects in later school years. These correlations might be even more negative for items that elicit a comparison only between math and language domains (Dickhäuser, 2005; Müller-Kalthoff, Helm, et al., 2017; Wolff, 2020; Wolff, Helm, Zimmermann, et al., 2018). However, our hypothesis would still be supported. Therefore, we believe that our conclusions are likely to hold true across these different operationalizations of internally referenced items.

Due to the design of this study, there are several other limitations. For example, we are not able to test whether the order of internally and externally referenced items might influence the results. To fully control for potential order effects, future studies would need to vary the presentation order within each measurement point. The current design also does not allow us to rule out the effect of changing from reading self-concept to English self-concept when students became older on the results. Nevertheless, our finding is aligned with studies that consistently used "reading" (Gunderson et al., 2017) or "German" (Wolff et al., 2020) as the verbal domain for all age groups. These studies reported a decline in the correlation between math and verbal self-concepts from elementary to secondary school using non-referenced items.

Finally, the sample in our data is primarily white and lower-middle to middle class. More longitudinal studies with more diverse samples are thus needed. The data we used in the analyses were collected between 1987 and 1996. Developmental patterns may have changed across different time periods, although we do not have a strong theory for the nature or directions of such changes.

CONCLUSIONS

Our study is the first to chart the changes in students' use of dimensional comparisons to form their ability self-concepts in Grades 1 through 12 by explicitly asking students to make self-evaluations based on dimensional comparisons. Our findings also suggest that the increase in students' use of dimensional comparisons largely drives the increasing differentiation of students' math and verbal self-concepts over time. Students' use of social comparisons shows a different developmental pattern that is unlikely to drive the differentiation processes in students' ability self-concepts. Finally, researchers should carefully consider the most appropriate wording for their assessments of students' ability self-concepts, as internally versus externally referenced items are likely to capture different developmental processes.

ACKNOWLEDGEMENTS

The research reported here was supported by the National Institute of Child Health and Human Development Grant HD17553 to Jacquelynne S. Eccles, Allan Wigfield, P. Blumenfeld, and R. Harold, and two grants from the National Science Foundation awarded to Jacquelynne S. Eccles (DRL-1108778 and HRD-1231347). The research was also supported in part by a Dissertation Grant to Sirui Wan from the American Psychological Association. D. H. Bailey was supported by a Jacobs Foundation Research Fellowship. The authors thank Nadia Chernyak, Barbara Sarnecka, and Sandra Simpkins for helpful comments on this project. The data necessary to reproduce the analyses presented here are publicly accessible at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/T1QCQQ. Code is available from the first author. The analyses presented here were not preregistered.

ORCID

Sirui Wan https://orcid.org/0000-0002-8750-0977

REFERENCES

Albert, S. (1977). Temporal comparison theory. *Psychological Review*, *84*(6), 485–503. https://doi.org/10.1037/0033-295X.84.6.485

Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science: A call to action. *European Journal of Psychological Assessment*, *38*(1), 1–5. https://doi.org/10.1027/1015-5759/a000699

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach.

Psychological Bulletin, *103*(3), 411–423. https://doi.org/10.1037/0033-2909.103.3.411

Beymer, P. N., Ferland, M., & Flake, J. K. (2021). Validity evidence for a short scale of college students' perceptions of cost. *Current Psychology*. https://doi.org/10.1007/s12144-020-01218-w

Bong, M. (1998). Tests of the internal/external frames of reference model with subject-specific academic self-efficacy and frame-specific academic self-concepts. *Journal of Educational Psychology*, *90*(1), 102–110. https://doi.org/10.1037/0022-0663.90.1.102

Cai, D., Viljaranta, J., & Georgiou, G. K. (2018). Direct and indirect effects of self-concept of ability on math skills. *Learning and Individual Differences*, *61*, 51–58. https://doi.org/10.1016/j.lindif.2017.11.009

Cimpian, A. (2017). Early reasoning about competence is not irrationally optimistic, nor does it stem from inadequate cognitive representations. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation (2nd ed.): Theory and application* (pp. 387–407). Guilford Press.

Dickhäuser, O. (2005). A fresh look: Testing the internal/external frame of reference model with frame-specific academic self-concepts. *Educational Research*, *47*(3), 279–290. https://doi.org/10.1080/00131880500287211

Dweck, C. S. (2002). The development of ability conceptions. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 57–88). Academic Press. https://doi.org/10.1016/B978-012750053-9/50005-X

Eccles, J. S., Arberton, A., Buchanan, C. M., Janis, J., Flanagan, C., Harold, R., Mac Iver, D., Midgley, C., Reuman, D., & Wigfield, A. (1993). School and family effects on the ontogeny of children's interests, self-perceptions, and activity choice. In J. Jacobs (Ed.), *Nebraska symposium on motivation, 1992: Developmental perspectives on motivation* (pp. 145–208). University of Nebraska Press.

Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, *21*(1), 225–241. https://doi.org/10.1111/j.1532-7795.2010.00725.x

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Ehm, J. H., Lindberg, S., & Hasselhorn, M. (2014). Reading, writing, and math self-concept in elementary school children: Influence of dimensional comparison processes. *European Journal of Psychology of Education*, *29*(2), 277–294. https://doi.org/10.1007/s10212-013-0198-x

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140. https://doi.org/10.1177/001872675400700202

Gaspard, H., Lauermann, F., Rose, N., Wigfield, A., & Eccles, J. S. (2020). Cross-domain trajectories of Students' ability self-concepts and intrinsic values in math and language arts. *Child Development*, *91*(5), 1800–1818. https://doi.org/10.1111/cdev.13343

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., Fischbach, A., & Preckel, F. (2014). "My questionnaire is too long!" the assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology*, *39*(3), 188–205. https://doi.org/10.1016/j.cedpsych.2014.04.002

Gunderson, E. A., Hamdan, N., Sorhagen, N. S., & D'Esterre, A. P. (2017). Who needs innate ability to succeed in math and literacy? Academic-domain-specific theories of intelligence about peers versus adults. *Developmental Psychology*, *53*(6), 1188–1205. https://doi.org/10.1037/dev0000282

Harter, S. (2012). *The construction of the self: Developmental and sociocultural foundations* (2nd ed.). Guilford Press.

Higgins, E., & Eccles-Parsons, J. (1983). Social cognition and the social life of the child: Stages as subcultures. In E. Higgins, D. Ruble, & W. Hartup (Eds.), *Social cognition and social development* (pp. 15–62). Cambridge University Press.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. Guilford Press.

Lauermann, F., Meißner, A., & Steinmayr, R. (2020). Relative importance of intelligence and ability self-concept in predicting test performance and school grades in the math and language arts domains. *Journal of Educational Psychology*, *112*(2), 364–383. https://doi.org/10.1037/edu0000377

Lauermann, F., Tsai, Y.-M., & Eccles, J. S. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy–value theory of achievement-related behaviors. *Developmental Psychology*, *53*(8), 1540–1559. https://doi.org/10.1037/dev0000367

Lohbeck, A., & Möller, J. (2017). Social and dimensional comparison effects on math and reading self-concepts of elementary school children. *Learning and Individual Differences*, *54*, 73–81. https://doi.org/10.1016/j.lindif.2017.01.013

Mac Iver, D. (1987). Classroom factors and student characteristics predicting students' use of achievement standards during ability self-assessment. *Child Development*, *58*(5), 1258–1271. https://doi.org/10.2307/1130619

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, *23*(1), 129–149. https://doi.org/10.3102/00028312023001129

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology*, *111*(2), 331–353. https://doi.org/10.1037/edu0000281

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. https://doi.org/10.1111/j.1467-8624.2005.00853.x

Marsh, H. W., & Yeung, A. S. (2001). An extension of the internal/external frame of reference model: A response to Bong (1998). *Multivariate Behavioral Research*, *36*(3), 389–420. https://doi.org/10.1207/S15327906389-420

Möller, J., & Husemann, N. (2006). Internal comparisons in everyday life. *Journal of Educational Psychology*, *98*(2), 342–353. https://doi.org/10.1037/0022-0663.98.2.342

Möller, J., & Köller, O. (2001). Dimensional comparisons: An experimental approach to the internal/external frame of reference model. *Journal of Educational Psychology*, *93*(4), 826–835. https://doi.org/10.1037/0022-0663.93.4.826

Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, *120*(3), 544–560. https://doi.org/10.1037/a0032459

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, *79*(3), 1129–1167. https://doi.org/10.3102/0034654309337522

Möller, J., Retelsdorf, J., Köller, O., & Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: An integration of models of relations between academic achievement and self-concept. *American Educational Research Journal*, *48*(6), 1315–1346. https://doi.org/10.3102/0002831211419649

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, *90*(3), 376–419. https://doi.org/10.3102/0034654320919354

Müller-Kalthoff, H., Helm, F., & Möller, J. (2017). The big three of comparative judgment: On the effects of social, temporal, and dimensional comparisons on academic self-concept. *Social Psychology of Education*, *20*(4), 849–873. https://doi.org/10.1007/s11218-017-9395-9

Müller-Kalthoff, H., Jansen, M., Schiefer, I. M., Helm, F., Nagy, N., & Möller, J. (2017). A double-edged sword? On the benefit, detriment, and net effect of dimensional comparison on self-concept. *Journal of Educational Psychology*, *109*(7), 1029–1047. https://doi.org/10.1037/edu0000171

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Development*, *49*, 800–814. https://doi.org/10.2307/1128250

Parsons, J. E., & Ruble, D. N. (1977). The development of achievement-related expectancies. *Child Development*, *48*, 1075–1079. https://doi.org/10.2307/1128364

Pohlmann, B., & Möller, J. (2009). On the benefit of dimensional comparisons. *Journal of Educational Psychology*, *101*(1), 248–258. https://doi.org/10.1037/a0013151

Schneider, R., & Sparfeldt, J. R. (2020). Academic competence and affect self-concepts in elementary school students: Social and dimensional comparisons. *Social Psychology of Education*, *23*(1), 233–257. https://doi.org/10.1007/s11218-019-09532-3

Sewasew, D., & Schroeders, U. (2019). The developmental interplay of academic self-concept and achievement within and across domains among primary school students. *Contemporary Educational Psychology*, *58*, 204–212. https://doi.org/10.1016/j.cedpsych.2019.03.009

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology*, *42*(1), 70–83. https://doi.org/10.1037/0012-1649.42.1.70

Skaalvik, E. M., & Skaalvik, S. (2002). Internal and external frames of reference for academic self-concept. *Educational Psychologist*, *37*(4), 233–244. https://doi.org/10.1207/S15326985EP3704_3

Skaalvik, E. M., & Skaalvik, S. (2004). Self-concept and self-efficacy: A test of the internal/external frame of reference model and predictions of subsequent motivation and achievement. *Psychological Reports*, *95*(3_suppl), 1187–1202. https://doi.org/10.2466/pr0.94.2.619-632

Stipek, D., & Mac Iver, D. (1989). Developmental change in Children's assessment of intellectual competence. *Child Development*, *60*(3), 521–538. https://doi.org/10.2307/1130719

Strickhouser, J. E., & Zell, E. (2015). Self-evaluative effects of dimensional and social comparison. *Journal of Experimental Social Psychology*, *59*, 60–66. https://doi.org/10.1016/j.jesp.2015.03.001

Umarji, O., McPartlan, P., & Eccles, J. (2018). Patterns of math and English self-concepts as motivation for college major selection. *Contemporary Educational Psychology*, *53*, 146–158. https://doi.org/10.1016/j.cedpsych.2018.03.004

Wan, S., Lauermann, F., Bailey, D. H., & Eccles, J. S. (2021). When do students begin to think that one has to be either a "math person" or a "language person"? A meta-analytic review. *Psychological Bulletin*, *147*(9), 867–889. https://doi.org/10.1037/bul0000340

Weidinger, A. F., Steinmayr, R., & Spinath, B. (2019). Ability self-concept formation in elementary school: No dimensional comparison effects across time. *Developmental Psychology*, *55*(5), 1005–1018. https://doi.org/10.1037/dev0000695

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wolff, F. (2020). Measuring comparison effects: A critical view on the internal/external frame of reference model. *Educational Measurement: Issues and Practice*. https://doi.org/10.1111/emip.12414

Wolff, F., Helm, F., & Möller, J. (2018). Testing the dimensional comparison theory: When do students prefer dimensional comparisons to social and temporal comparisons? *Social Psychology of Education*, *21*(4), 875–895. https://doi.org/10.1007/s11218-018-9441-2

Wolff, F., Helm, F., Zimmermann, F., Nagy, G., & Möller, J. (2018). On the effects of social, temporal, and dimensional comparisons on academic self-concept. *Journal of Educational Psychology*, *110*(7), 1005–1025. https://doi.org/10.1037/edu0000248

Wolff, F., Sticca, F., Niepel, C., Götz, T., Van Damme, J., & Möller, J. (2020). The reciprocal 2I/E model: An investigation of mutual relations between achievement and self-concept levels and changes in the math and verbal domain across three countries. *Journal of Educational Psychology.*, *113*, 1529–1549. https://doi.org/10.1037/edu0000632

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wan, S., Lauermann, F., Bailey, D. H., & Eccles, J. S. (2023). Developmental changes in students' use of dimensional comparisons to form ability self-concepts in math and verbal domains. *Child Development*, *94*, 272–287. https://doi.org/10.1111/cdev.13856