

Spaced Repetition for Schools: the White Paper

(DRAFT—DO NOT CIRCULATE YET)

I. Introduction

The advent of computers has made spaced repetition available to the masses in the form of spaced-repetition software (or SRS), and provided a wealth of information on their efficacy. The classic [article/blogpost by Gwern](#) includes a wide swath of citations to studies (both formal and informal) attesting to the usefulness of SRS, and will serve (for our purposes) as a literature review on the subject. SRS is used by tens of thousands of people for language-learning, mathematics, science and trivia. It is particularly heavily used by medical students, whose chosen profession demands the absorption of a punishingly large amount of content knowledge. The most commonly-used SRS platform is the open-source [Anki](#), but alternatives such as [Mnemosyne](#) and [SuperMemo](#) are also reasonably popular.

However, despite its effectiveness, SRS remains little-used outside of a handful of niche contexts such as med school. I will argue that this is due primarily due to the deficiencies of the existing platforms, and that the time has come for an SRS system built for classroom usage rather than the disciplined autodidact. In a former stint as a Latin teacher, I attempted to include Anki in the classroom and encouraged students to use it, but soon ran up against a number of serious issues:

- The core of Anki's codebase dates to the early-to-mid-2000s, a time when software was generally downloaded and installed on the user's machine, and the user could be expected to configure and tweak its settings as needed. In the 2020s, most software is browser-side, and although Anki does allow the user to do reviews in the browser, the browser-side review platform is a mere ghost of the full program. **Most students today use Chromebooks onto which software cannot be installed locally and therefore cannot really use Anki**; this is the biggest of a number of high hurdles.
- Anki is extremely customizable and powerful for "power users", but the learning curve is steep, and the casual user is tempted to give up early on because the default

settings are mediocre at best. This is perhaps most acute in the field of card and template design. It is possible to design custom cards with automatic formatting or other unusual behavior, but doing so requires the knowledge of at least HTML and in most cases also Javascript. There is a great deal of low-hanging fruit in the UI which could be picked with only a few lines of extra code.

While Anki's card- and template-design abilities are powerful (considerably more so than those of competitors such as Mnemosyne), most of that power is inaccessible to casual users. Formatting options such as color must be selected manually and laboriously for every field, or implemented via custom scripts. Cloze deletions exist only for text. Language learners must switch between keyboard layouts every time they want to use diacritics, or copy-and-paste from a third-party keyboard. There is no native support for creating diagrams, which must be made in another application (such as Figma) and imported as images. The native math package, MathJax, allows virtually no additional formatting, which makes it difficult to use best-case practices when designing cards for mathematics, unless the user is willing to invest time and effort into writing unintuitive, tedious and kludgy LaTeX code. There is no native support for mass card manufacture via spreadsheet; users must create a CSV (comma-separated file), usually by converting and exporting from Excel.

- There is no support for complex relationships between cards other than sibling cards within the same template. Good cards mimic the structure of human memories, whose interrelations are much more complex than 'bury siblings until the following day'. Native support for this would make reviews both more effective, with much less friction during sessions.
- Aside from shared decks, Anki has no support for usage by groups of learners rather than individuals. The teacher cannot see whether or not students have been doing reviews, or how well they have done on them, except by checking their screens individually. Nor can a teacher tweak cards in students' decks after a shared deck has already been downloaded, make cards for students on an ongoing basis, or see which cards have been giving students trouble. As noted above, because Anki is

downloadable rather than browser-based, it cannot be used by most students in most classrooms anyways.

- Because Anki is open-source, non-profit software, there is very little incentive to change. Development is subsidized by iOS users, who must pay (\$24.99 at the time of writing) to download the app version of Anki. It is unclear how many downloads the iOS app version of Anki has garnered, but it is currently the #1 paid app in the 'education' section. It has clearly done quite well for its main developers, but by the same token is stuck in a local maximum.

The above deficits also apply, sometimes to an even greater degree, to Anki's main competitors (Mnemosyne and SuperMemo), so I will pass over the alternatives and use Anki as my point of departure. Anki is the best of a severely underoptimized batch, and the time has come for something new. This essay will dig into each of the above criticisms in turn and envision a competitor that does things differently.

II. Platform

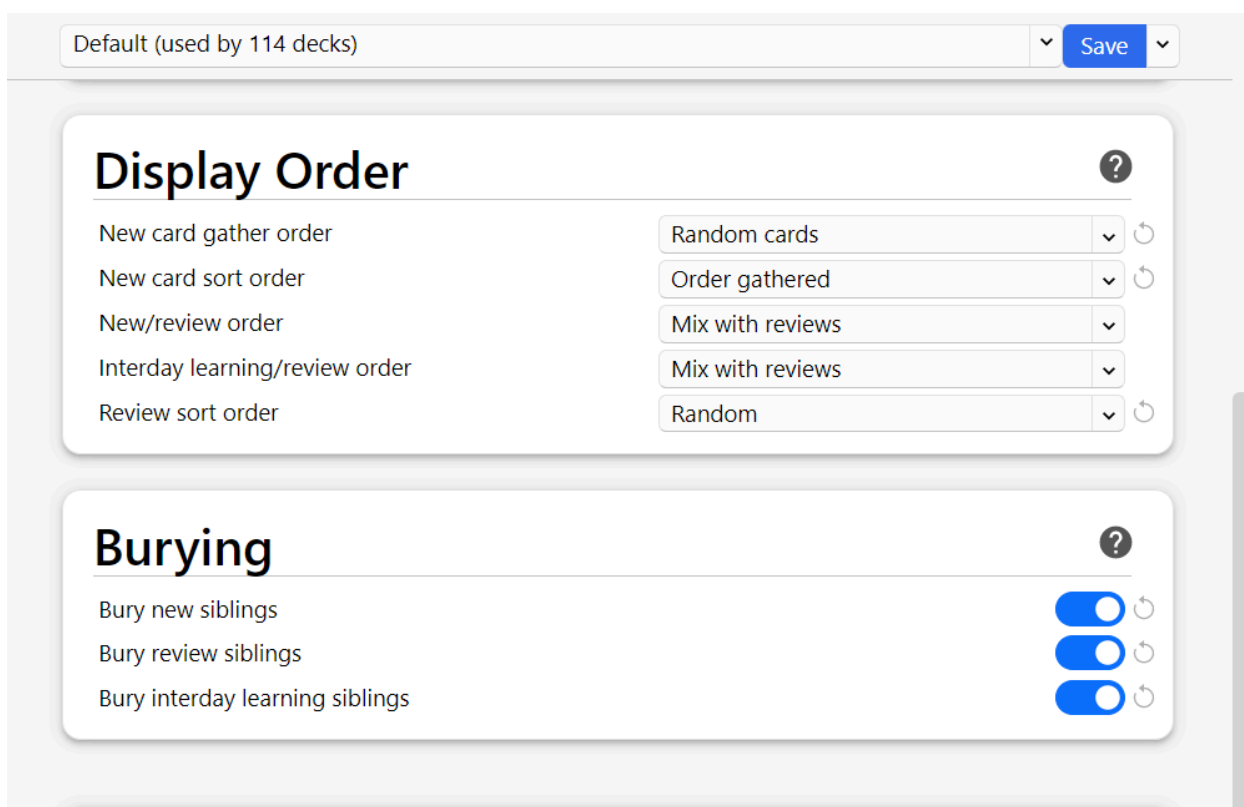
There is little to say here that has not been said. The software ecosystem of 2006, when Anki was released, was one of software distributed over the Internet or CD-ROM and installed locally on the user's computer. Nearly two decades on, most software is browser-side, and users often expect that software will be in-browser; students born since ~2010 or so are often bamboozled by the prospect of installing software locally. This alone would make a strong case for a native browser-side app.

More pressingly, however, most students in contemporary K-12 classrooms use Chromebooks, which cannot install software locally without carefully jailbreaking the device, which few students, parents or districts can or will want to do, not least because even a moderately experienced attempt is liable to brick the device. It is possible for Chromebook users to use AnkiDroid, the Android version of Anki, but—although AnkiDroid is superior to Anki's native web-based reviewer, it is still a far cry from the desktop version. The conclusion is therefore not merely strong but unassailable: **any SRS platform hoping to see wide**

classroom usage must be browser-based from day one, and none of the available options are.

III. *UI Design and the Learning Curve*

Anki's user interface is kludgy. Virtually no power users use the default settings, which suggests that the default settings are in dire need of optimization¹. Users wishing to improve their review experience are accosted with a bewildering array of dropdowns and buttons: what is 'insertion order'? What is the difference between 'card gather order' and 'card sort order', and why should I care about this? The tooltips explain these mechanics to some extent, but not why one might want to pick one option over another.

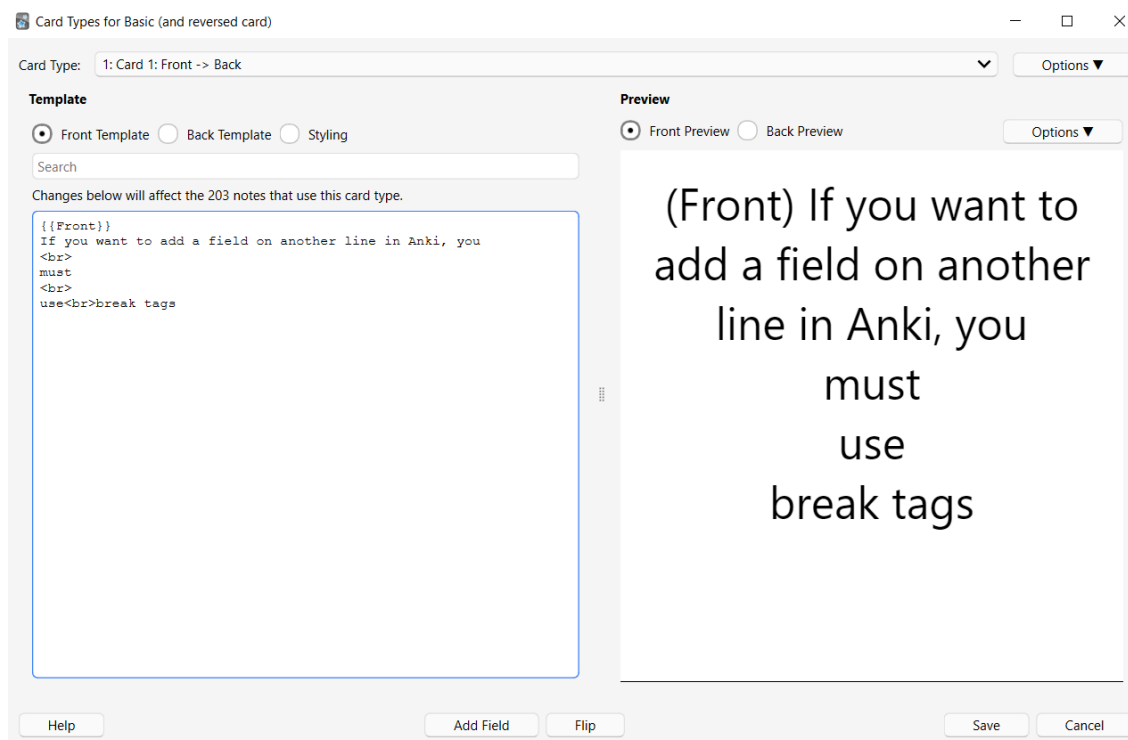


Most users therefore end up Googling for a set guide to improving their Anki settings and follow it blindly—in effect merely switching from one set of blind default settings to another.

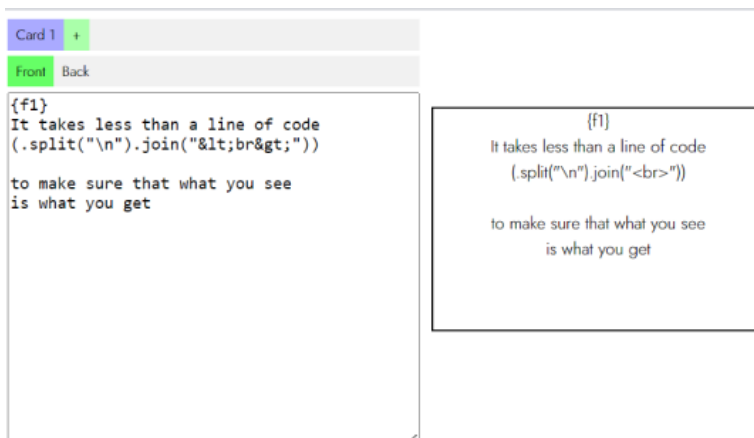
¹ However, the default settings have improved somewhat in the last few months since Anki switched to the FSRS algorithm. The FSRS algorithm is, happily, open-source, so we do not need to worry about exactly what the scheduling algorithm will look like—not, at least, until much later on.

A more serious problem is that of card and template design. There is wide agreement among serious SRS devotees that good card design is a make-or-break necessity for using spaced repetition effectively. Good cards fit comfortably within the contours of the user's cognition; bad cards create "cognitive friction" and make review sessions tedious and unpleasant. It is possible to make good cards with the default card templates, but for many purposes optimal cards require a customized template. Here Anki fails:

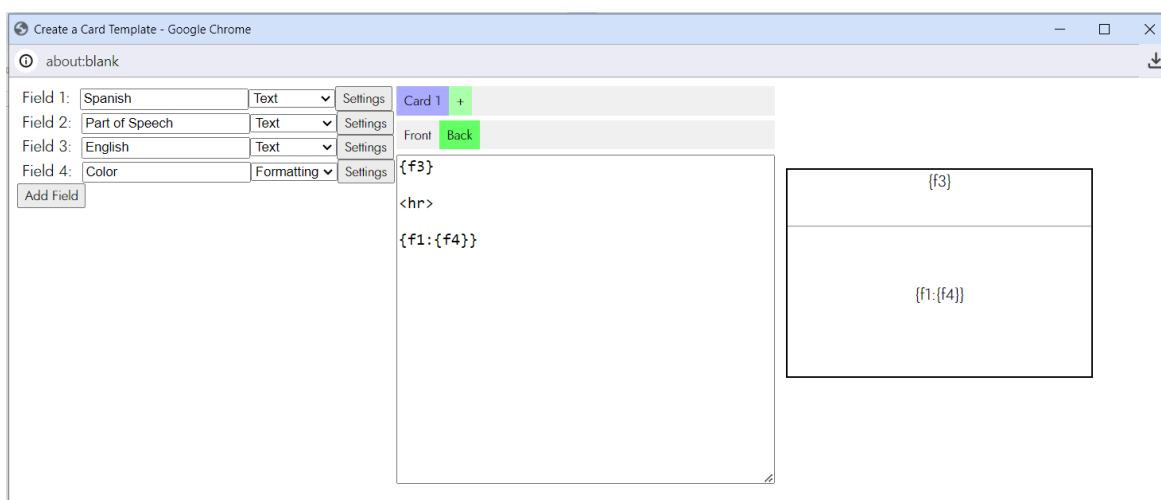
- The template-editor is raw HTML, not what-you-see-is-what-you-get (WYSIWYG):



But the basics of a WYSIWYGeditor are not difficult:



- In a similar vein, consider advanced formatting options such as color. It is well known that the human mind is very good at remembering color and that color-coding makes it much easier to remember arbitrary information such as gender, tone or the identity of an atom on a molecular diagram. In Anki, color-coding must be done manually via a color-picker which creates a new window, which creates friction in card design. This becomes tedious when making more than a few cards in a session. I eventually got around this in my own vocabulary cards by adding a custom field to my vocabulary cards which takes a letter for gender (*m*, *f*, *n* or *e*²) and, via custom Javascript, applies corresponding coloring or highlighting to the text of the word. But it was not at all obvious that this is possible; it did not occur to me until after a few years of heavy Anki usage. We would prefer to do something like this:



This has not yet been implemented in the course of my work coding the toy model—which has only been ongoing for a couple days old at the time of writing—but it is hopefully much more self-explanatory than an ugly pile of Javascript. Field 4, “Color”, will show up in the card editor as a dropdown or radio-button selector, and applies as formatting to Field 1 (Spanish) on the back side of every Card 1. Similarly, it should be possible to add basic formatting such as **boldness**, *italicization* or underlining to text via clicking a button, as in Microsoft Word, rather than using HTML tags³.

² For epicene nouns, e.g. Latin *sacerdōs* or Hebrew תַּהוֹם, which may be either masculine or feminine.

³ This is possible in Anki’s card-creator, but it cannot be done automatically in the card-template designer without using HTML.

This is not to say that custom Javascript should not be possible. Power users of a replacement for Anki should, generally, be able to do anything power users of Anki currently can. But it is **much** more important that casual-to-medium-intensity users of the replacement be able to do *most* things that Anki power users can, including automatic custom formatting and good card design, without feeling overwhelmed.

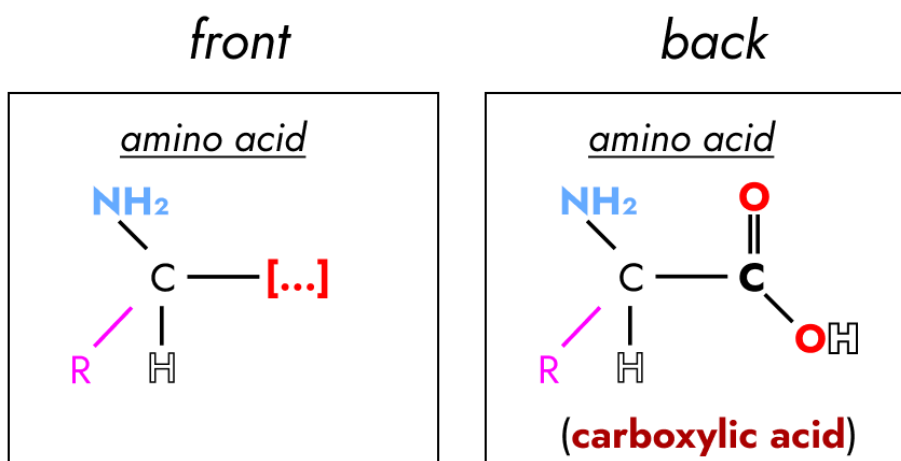
- Students of mathematics or physics—or any natural science where mathematical formulas are key—are also underserved by Anki. Anki's built-in MathJax library has no access to color-coding or other formatting; it allows cloze deletions, but these are black-and-white and thus difficult to see. More advanced math cards must be made with LaTeX, which adds yet another steep learning curve. A what-you-see-is-what-you-get math editor of the sort that WolframAlpha uses, with formatting options, is both feasible and desirable.
- Language-learners must switch keyboards, or copy and paste from a third-party program, to use diacritics or other alphabets; this, again, creates a great deal of friction in card manufacture. There is no reason that users should not be able to define or pick a custom text-replacement mapping that would apply to certain fields. We could imagine such a thing for German, for example, allowing us to write `<scho"n>` in the card-editor and have it convert automatically to *schön*.

The screenshot shows a web browser window titled "Add Character Replacement - Google Chrome". The address bar shows "about:blank". Below the address bar is a form with an "Add" button. The form contains several rows of input fields for mapping characters to German diacritics:

a"	→	ä
o"	→	ö
u"	→	ü
S"	→	ß
A"	→	Ä
O"	→	Ö
U"	→	Ü

As with automatic coloring, this can be done in Anki with Javascript, but doing so is tedious, especially for alphabets other than Latin. It should not require Javascript to convert <chislo> to <число>, <mantha'nw> to <μανθάνω>, or even <k'Atab> to <כְּתָב>⁴.

- Cloze deletions are powerful and efficient, but exist only for text; there is no built-in diagram maker or any other way to handle diagrams or pictures other than as entirely static image files. Students of chemistry or biology might want to be able to do something like this:



Anki has no capabilities for such a thing; two separate images must be designed in a third-party application such as Figma and copy-pasted into Anki's card maker. While there is no need to give an Anki-replacement the full powers of real image-editing software, a lightweight version for simple image-editing and diagram-making, built with common card formats in mind, is certainly feasible.

- Anki has no built-in way to manufacture cards *en masse*; users must use Excel and export a CSV file to do so. This is not a big ask for power users, but it represents yet another barrier to adoption by people who want things to Just Work—including most students and teachers.

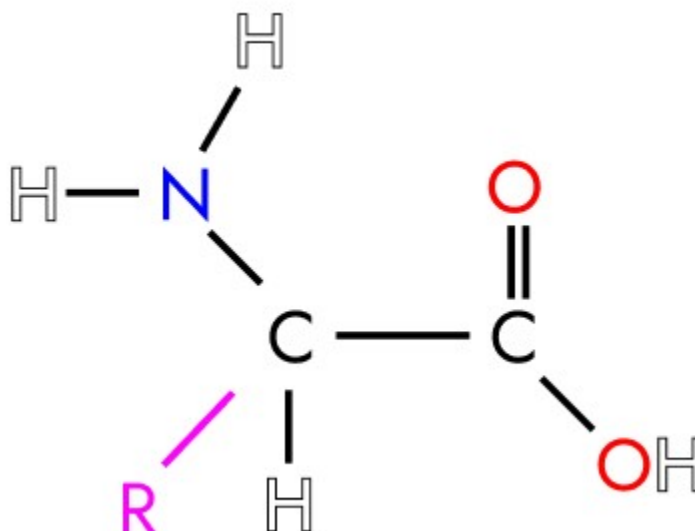
⁴ Chinese characters or Japanese kanji will not, of course, be quite so simple; but text-to-character libraries are numerous.

Every single one of these deficiencies is solvable, and some (such as automatic text replacement for diacritics or foreign alphabets) are very simple to solve. But Anki's developers have never thought to solve them, and Anki therefore remains a niche product.

This section would be incomplete without a discussion of LLMs and other “artificial intelligence” applications; many Anki users have used LLMs to great success in creating mnemonic images for things they need to learn. There is probably a place in a modern, classroom-based SRS platform for an LLM, though I am not yet ready to place any bets on what it might be. My larger point is that the deficiencies of existing SRS systems are entirely orthogonal to the use of LLMs. Nothing I have discussed in this section requires an LLM, and every proposal in this white paper could have been deployed fifteen years ago.

IV. *Card Relationships*

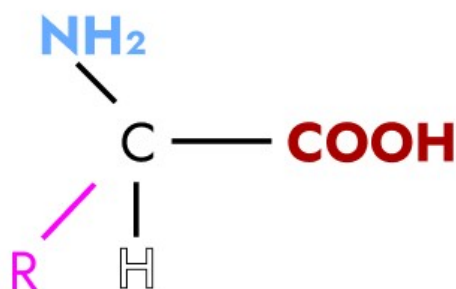
Consider, again, the structure of an amino acid, with atoms color-coded for convenience:



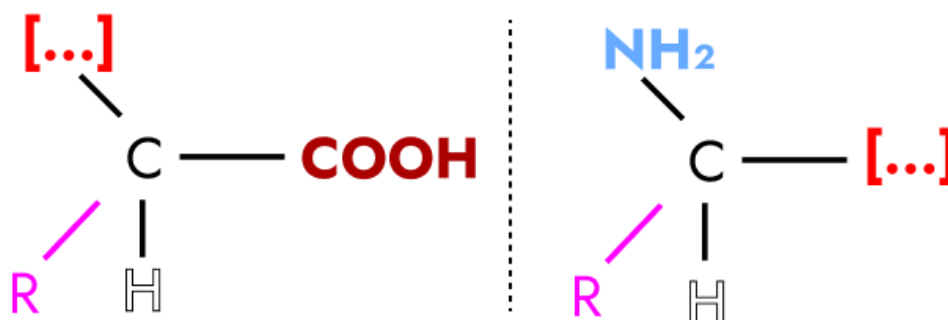
The general consensus among SRS enthusiasts is that ‘draw an amino acid from scratch’ is a bad prompt, because it’s insufficiently atomic. It demands too much information of the user at once, and should be broken up into parts.

I disagree. A biochemistry student may indeed be asked to draw the structure of an amino acid from scratch on a quiz; that's a perfectly valid prompt and one that we might well wish to review. What is true is that if the student tries to learn the structure of an amino acid from scratch as a whole unit, **they are likely to fail.**

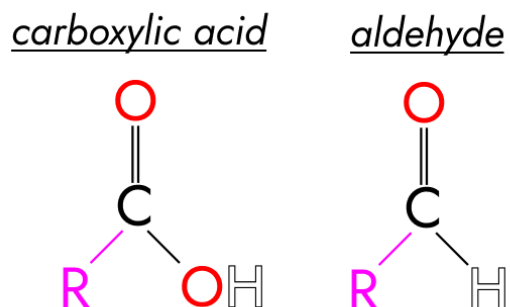
This is not because it is impossible to store the structure in one's head as a single chunk. Rather, packaging it into a single memory chunk is possible *only after its constituent sub-chunks have been learned.* An amino acid, like all organic molecules with more than a handful of atoms, is composed of multiple recognizable subcomponents. For an amino acid, the two recognizable functional groups are the amine to the "northwest" of the bonding carbon, and the carboxylic acid group to its "east".



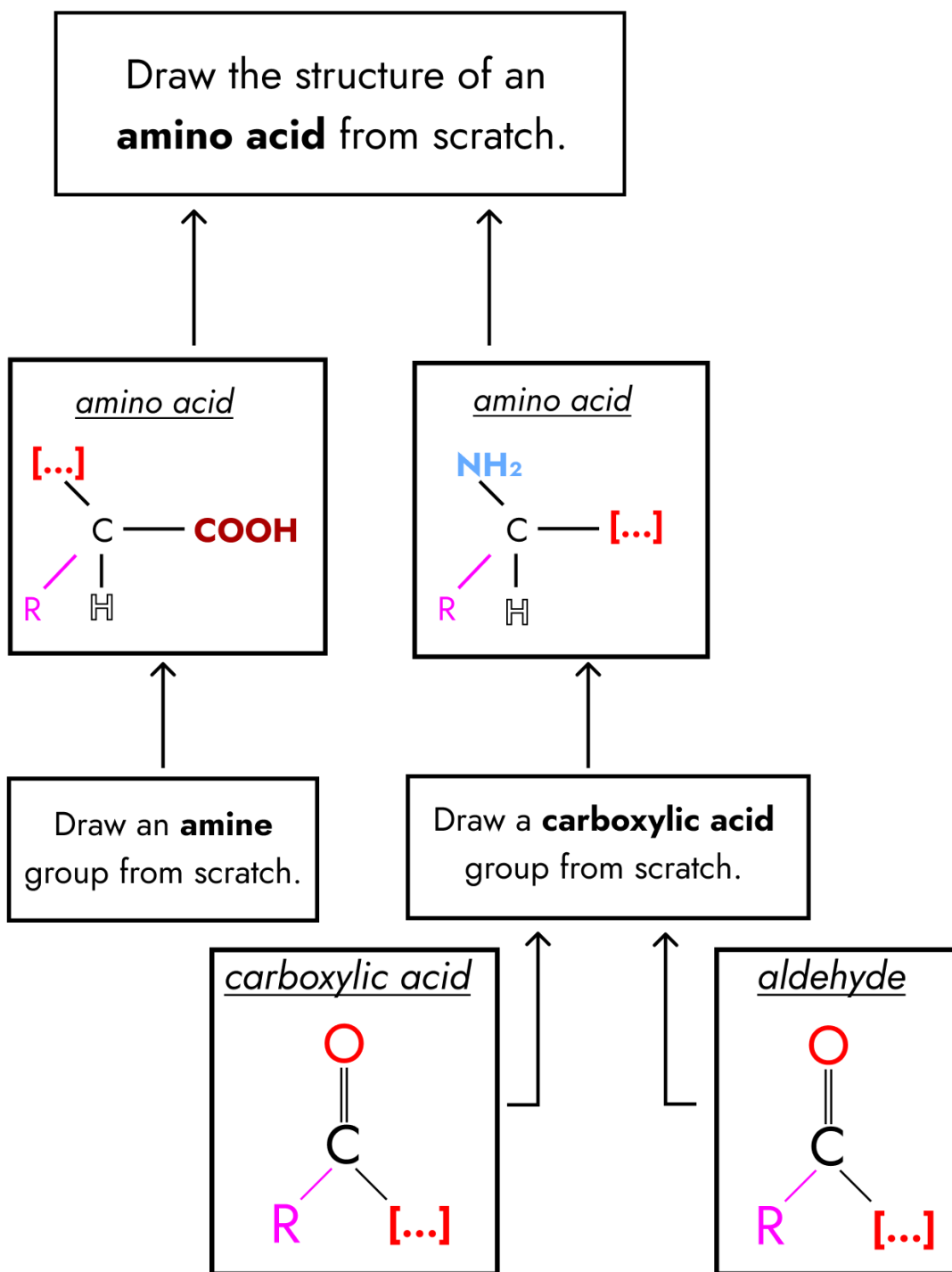
Solidifying the structure of an amino acid into a single chunk thus has prerequisites: we must know the functional groups. We can imagine the following cards to test that knowledge:



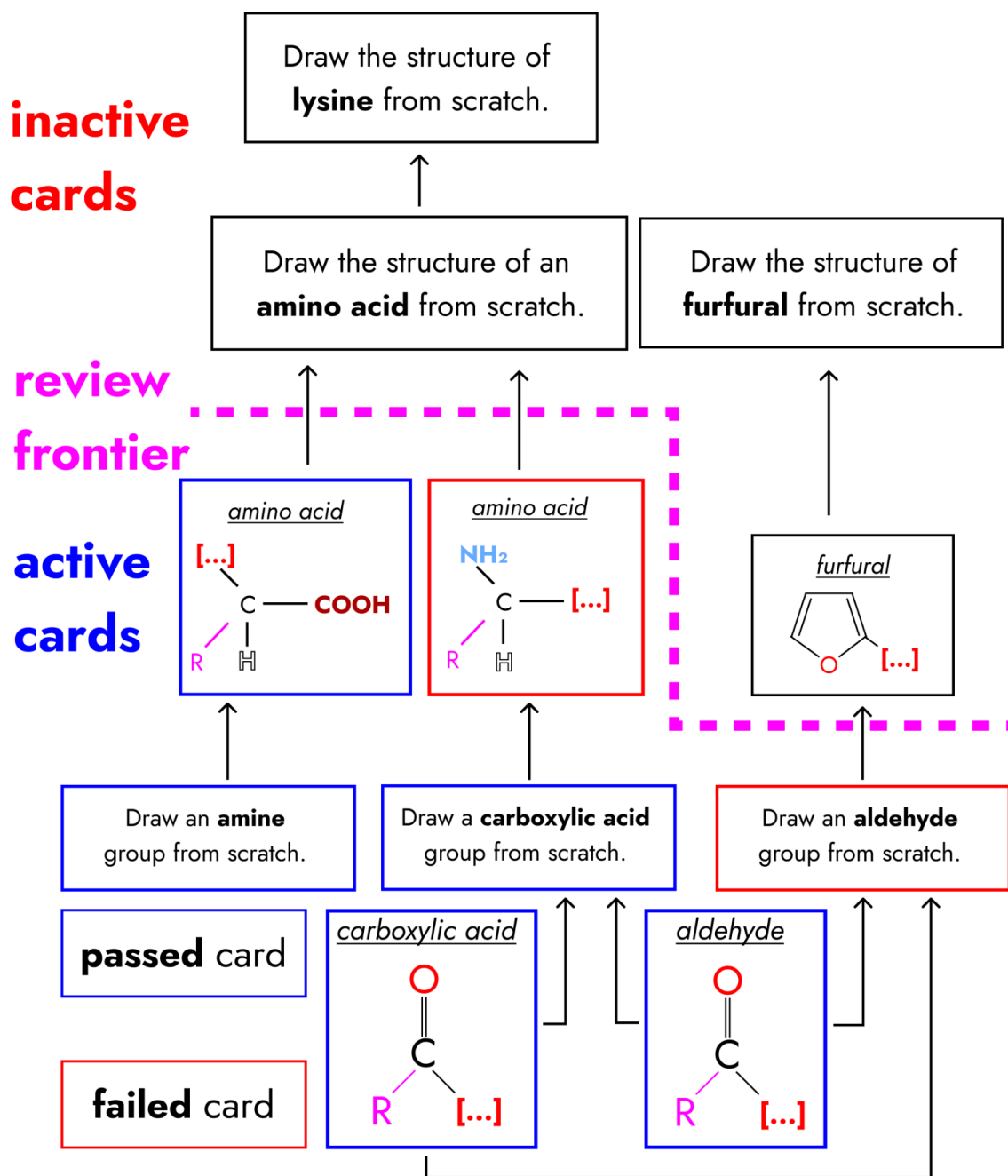
Even these might themselves have prerequisites. If a carboxylic acid group's hydroxide is swapped out for a hydrogen atom, the result is an aldehyde group:



The structure of an amino acid can certainly be stored in a single mental chunk, but not until its component chunks have been 'solidified'—a process which itself relies on those components' subcomponents. In the diagram on the next page, an arrow $A \rightarrow B$ means 'card A is a prerequisite for card B':



We probably don't want to test a student on the full structure of an amino acid until we're fairly sure they have a decent handle on its prerequisites—doing so will just create mental friction. There will be a *frontier* of active cards based on whether or not all their prerequisites have been, if not mastered, at least marked as 'correct' on the last review.



The extent of Anki's card-relationship abilities is "sibling burial". If a template produces more than one card—e.g. a vocabulary-card template which produces one English → Spanish card and one Spanish → English card—then those cards are considered 'siblings', and we can set Anki to 'bury' cards on days when one of their siblings is due for review. In effect, this tells Anki to show only one card from a template per day.

But ideally, we'd like a replacement for Anki to 'follow the mind' and do a better job of imitating and reinforcing memories as they are actually structured: as a **network** of chunks and concept rather than a collection of atomized facts. There is an add-on that allows this sort of hierarchical burial structure, but it only applies to cloze cards, and requires a certain amount of hacking Anki's review system. There is no way to do this for other kinds of cards or for cards belonging to different notes⁵, and even if you try to use the sibling-burial system to proxy it, you'd need to create a new template for every separate card-prerequisite hierarchy. The result is that the same chunk of information has to be added to, and reviewed in, cards from multiple templates, which means those cards interfere with each others' memory curves. We can do better.

V. *Spaced Repetition for the Classroom*

I used to teach middle-school Latin, and tried getting my students using Anki. A few took to it like ducks to water and, as far as I know, are still using it to great success. Most didn't, for a number of reasons:

- As noted in section II, many of my students used Chromebooks, which meant Anki was inaccessible to them.
- Most schoolchildren make poor autodidacts. They are bad at recognizing what they're having trouble with and even worse at creating cards or study materials to hone in on that issue. For example, a student might have trouble distinguishing *dūcō* 'I lead' and *dīcō* 'I say'; they only differ in a single vowel, so the wires are easily

⁵ A note is an instantiation of a template. Thus, in a vocabulary deck, *perro* → *dog* and *dog* → *perro* will be different cards, but belong to the same note (making them siblings).

crossed. A cloze deletion of the form **d[...]*cō*** is a good way to review this, but knowing that you need to do that requires a level of metacognition most students don't develop until at least high school. Teachers need to be able to do that for them...

- ...but the only way for teachers to do so is to make shared decks for download. There is a consensus among Anki devotees that shared decks mostly work well for 'prepackaged' bodies of knowledge such as medical school courses; I've never been able to get anywhere with premade vocabulary decks (and it's probably not coincidental that a medical student's metacognition abilities are far better than the average eighth-grader's). The only way to edit or tweak a shared deck after it's been downloaded is to ask students to download it again, which destroys their progress. Ideally, a teacher should be able to update, add or delete cards from students' decks as needed.
- Even if it were possible—in, say, a school where all the students are given a Macbook—to get everyone using Anki, students' progress must be checked manually, student by student. There is no easy way to see what cards are giving students trouble, or whether students have been doing reviews at all.
- Autodidacts can generally be trusted to be honest with their reviews. Students can't be; left to their own devices, they will just press 'easy' on any card they don't want to deal with (I would certainly have done this in high school if there were no checks). You could make all cards type-in-the-answer, but my experience has been that this creates friction and makes the experience unpleasant.

A potential alternative is to check students on a *selection* of reviews using 'quiz mode'. At the end of a session, our SRS system can pull a handful of cards from that day's review and the most recent prior reviews that the student claims they passed, and ask them to type in the answer with a time limit⁶. Incorrect answers, or answers requiring the student to draw a diagram or speak into a microphone, can be sent to

⁶ The time limit could, of course, be adjusted for students with extended-time accommodations.

the teacher anonymously to be quickly graded (perhaps the student made a typo). Incorrect answers will be marked as a fail on the associated card; at the end of the quarter the teacher can pull up quiz-mode results and add them to the gradebook.

- Anki is not-for-profit, though as noted in the introduction, it seems to be something of a golden goose for its top developers. But a good SRS system designed for classroom usage is something schools will pay for.

I should emphasize that this does not mean that the design of such a system needs to revolve around the classroom to the exclusion of individual usage. As noted, there are a number of simple possible improvements to Anki's design that would greatly enhance the individual experience as well. But, at this point in time, there is no SRS system on the market—open-source or otherwise—which is built for classroom usage by default⁷. Contemporary education is beset by the *bêtes noires* of learning loss and subpar content knowledge; compared to their counterparts twenty or even ten years ago, today's students are learning less at school, and forgetting more of what they do learn. Spaced repetition is not a panacea for these (no single platform or application can be), but it *does* constitute an effective tool that schools and teachers can and should be using—if that tool works out-of-the-box in a classroom context, is designed to minimize friction for teachers and students without sacrificing customizability, and is built for the browser.

⁷ I should note that in the past few years there have been tentative steps in this direction by a handful of established educational technology companies. Quizlet, for example, has recently introduced a spaced-repetition mode. But these are generally afterthoughts. Quizlet's structure is built for creating, and reviewing, a single deck of cards in one session for a specific unit—polyatomic ions, for example, or the vocabulary from Chapter 6 of the French textbook. This **does not** make Quizlet a true spaced-repetition system; the real power of Anki or Mnemosyne derives from the fact that a user may have four or even five figures' worth of cards in a single deck, but will only need to review a tiny fraction of that number on a given day.

VI. *Prospects and Challenges*

Edtech is a notoriously difficult sector to do well in; while public school districts have large budgets, much of which they waste by any halfway objective measure, their willingness to try something new is constrained by strict procurement rules. Early on, the target market is thus likely to be private schools and homeschooling pods (whose procurement processes are much simpler), as well as autodidacts (who have no procurement processes beyond their own wallets).

It should be remembered that there are many levels of potential success. American K-12 education is a gigantic market, with 54 million students, about 9% of whom go to private schools. A browser-based, classroom-oriented SRS platform could still do very well even if it only captures a small fraction of that market. Suppose, for example, that the market share were to level out at twenty-five thousand students' worth of private schools, each paying \$50 per student per year (very much on the low end for educational-technology subscriptions). Such a company would be seeing revenues of \$1.25 million yearly; assuming a 40% profit margin and a price-to-earnings ratio of 20, such a company would be valued at \$10M. It is probably not too delusional to say that—contingent on the company successfully bringing its products to market and finding customers in the first place—these figures might well turn out to be underestimates.

Building a platform of this type will certainly be challenging. But it is feasible, and the time is now ripe to do so. There is little competition and a large, well-funded market facing serious challenges a modern SRS system can offer solutions to. It's time to build one.

Campbell Nilsen

Ashburn, VA

February 22nd, 2024

campbell.nilsen@gmail.com