Understanding and Avoiding Plagiarism

Below is the key information contained in the module on understanding and avoiding plagiarism created for Caltech undergraduates.

SECTION 1: WHAT IS PLAGIARISM?

Caltech's undergraduate handbook defines plagiarism as "the appropriation of another person's ideas, processes, results, or words without giving appropriate credit."

This definition of plagiarism is also used in Caltech's Policy on Research Misconduct, which applies to all faculty and researchers on campus. The National Science Foundation, National Institutes of Health, and other major U.S. research organizations use the same or very similar definitions.

Our community and the broader scientific research community have agreed that when you use another person's ideas, processes, results, or words in your writing, you have to clearly let a reader know that you are doing so. We utilize academic citation style systems to give appropriate credit.

Why is avoiding plagiarism so important to our community?

The core of Caltech's mission is to "expand human knowledge." Such expansion involves the contribution of new ideas, processes, results, and words. However, new contributions necessarily build on the prior work of others. Academic writers must distinguish for their readers between which parts of their writing are original and which parts come from other writers. This practice of attributing sources allows academic readers to understand the nature of each writer's contributions. It also gives fair credit to the other thinkers on whose work our own work builds.

All academic writers, including students, at Caltech are expected to give credit in this manner. When you are writing in the context of a course, giving appropriate credit allows your professor to see how you are making use of existing ideas, processes, results, and words, as well as what new things you are contributing. Making this distinction clear for your professor allows them to assess whether or not you are meeting their learning goals for the course.

It is important to note that the beliefs that underlay the expectations U.S. academic writers have for working with texts are connected to cultural values such as intellectual property and individualism. These values are not universal. In other cultural contexts, writers may have different practices for acknowledging sources. What is described in this module is a set of beliefs and practices that shape academic writing in the U.S. university and which also shape international scholarly publishing.

What happens when an undergraduate writer at Caltech is suspected of plagiarism?

If an undergraduate writer submits a paper or other text in a course and the instructor believes it contains plagiarism, the instructor will give the text to our Board of Control, the body that enforces our academic honor code. The Board of Control investigates the facts and makes recommendations to the Undergraduate Dean, who then decides whether the honor code has been violated in academic contexts and what the results of a violation should be. The process aims to educate the student about academic integrity and to protect the Caltech community from further violations. Potential results of an honor code violation stemming from plagiarism include failing the assignment, failing the course, and/or being placed on leave from the Institute to allow time to

consider and plan future compliance with the honor code. You will learn more about our honor code during orientation.

Caltech's honor code is central to how our learning and research community works. It states, "No member of the Caltech community shall take unfair advantage of any other member of the Caltech community." This honor code provides our community members with valued freedoms and assurance of living and learning in an ethical, supportive community.

Plagiarism in student writing violates the honor code by taking unfair advantage of other student writers who properly attribute credit to themselves and others, and by taking unfair advantage of course instructors who reasonably expect to be able to distinguish between original student work and ideas, processes, data, and words drawn from other sources.

SECTION 2: TYPES OF PLAGIARISM AND HOW TO AVOID THEM

There are a number of ways a writer can fail to give appropriate credit for another writer's ideas, processes, results, or words. In the following section, we will detail seven types of plagiarism that can occur in academic writing. We will also provide guidance for how to avoid each type of plagiarism. Here are the types of plagiarism we will cover: theft; outsourcing and overcollaboration; patchwriting; self-plagiarism (also called text recycling); inadvertent appropriation; citation errors; and common knowledge problems.

Plagiarism Type 1: Theft

Definition: This type of plagiarism occurs when a writer knowingly appropriates another writer's ideas, processes, results, or words and purposely provides no attribution or misleading attribution. In such cases, a writer intends to present someone else's work as their own.

Such actions clearly constitute plagiarism and violate the values of our academic community.

How to Avoid: Writers typically resort to deceptive writing practices like theft under stress. Most often, this happens if a writer has not planned sufficient time to compose a text on their own. The writer may begin to feel desperate, because they have created a situation in which they have to compose a text in a short amount of time.

There are two options for avoiding theft. The first and best strategy is to start writing projects early, so that you have plenty of time to assess how difficult they will be and how much time they will take. This also allows plenty of time to draw on your professor, TAs, and Caltech's Hixon Writing Center to support your writing. Do not put yourself in a stressful position where theft seems tempting.

If you fail to leave yourself sufficient time to write and seek support, the second strategy is to reach out to your professor (via email or in person) and let them know that you will not meet the writing deadline. Ask what options the professor can offer you, if any, to still get some credit for the assignment if it is submitted late. It is far better, in both the short and long term, to earn reduced credit or to fail an assignment than to submit writing that contains theft and thereby violates the honor code.

Plagiarism Type 2: Outsourcing and Overcollaboration

Definition: Outsourcing occurs when a writer asks someone else to compose part or all of a text for them.

It is common for academic writers to collaborate on projects. If you examine academic journals, many articles have multiple authors who had different roles in producing the manuscript. In courses, you may sometimes be asked to work with other students to write a paper together or to jointly compose other texts like posters or slides.

However, academic readers expect that all contributing authors are clearly and explicitly acknowledged as authors. It is a form of plagiarism to have an unacknowledged author who produced ideas, processes, results, or words in a text.

A problem related to outsourcing is overcollaboration. Many courses at Caltech encourage collaborative learning, and this includes encouraging you to seek feedback on your work as a writer. Seeking and responding to such feedback is a key way we get better at writing. In general, most professors will encourage you to share written drafts with peers and staff in the Hixon Writing Center. In such conversations, a reader will help you understand the strengths and limits of your work. But when it comes to actually revising and editing your draft, it is up to you, the author, to make those changes yourself.

How to Avoid: As with theft, outsourcing typically happens when a writer is trying to cope with a stressful writing situation. The writer may feel they lack the time to compose a text on their own, or they might feel they lack the knowledge or skill to compose the kind of text they are being asked to write. The writer might decide to pay someone to write a paper for them, or to ask a parent or friend to help them write a paper. This kind of academic dishonesty is unacceptable. It both impedes and misrepresents a student's learning.

As we previously advised, time management can be key to avoiding this kind of extreme stress. In cases where the temptation to outsource comes from insecurity about your ability to rise to a writing challenge, you need to reach out for help. College involves navigating many difficult assignments, and your professors expect you to ask for help when you need it. It is not a sign of weakness to seek out support during a professor's office hours or from the professional and student staff in the Hixon Writing Center. Seeking such help is often a crucial part of the learning you will do in college.

As for avoiding overcollaboration, check course collaboration policies to see what kinds of support you are allowed to seek from peers or other readers of your written work in a particular course. If a course collaboration policy does not address writing, ask your instructor for guidance.

Plagiarism Type 3: Patchwriting

Definition: The term patchwriting may not be familiar to you. Specialists in academic writing coined this term to describe a writing practice that involves "copying from a source text and then deleting some words, altering grammatical structures, or plugging in one-for-one synonym-substitutes" (Howard, 1992, p. 233). Another way of thinking about patchwriting is that it is a paraphrase that is too close to the original. Patchwriting is considered to be a form of plagiarism, even if you let the reader know through citation that you are paraphrasing another writer's work.

Some of you may be surprised to learn this. Novice academic writers are sometimes taught to use patchwriting as a writing strategy. When young writers are first working with the texts of other writers, they may have trouble understanding those texts and putting them into their own words in an authentic way. However, the practice of copying a text and then altering it until it seems different enough to be a fair paraphrase is not an accepted academic writing practice. As a college writer, you should never patchwrite.

Reference:
Howard, R. M. (1992). "A plagiarism pentimento." *Journal of Teaching Writing*, *11*(2), 233-246.

How to Avoid: Paraphrasing a text—putting another writer's key ideas in your own words—is challenging. This is because to properly paraphrase a text, you have to deeply understand it. The best way to avoid patchwriting is to follow two guidelines.

First, only paraphrase texts you have actually read and understood. If you haven't spent enough time with a text to understand it, you haven't yet earned the right to paraphrase it in your own work.

Second, when you do paraphrase a text, set aside the book, PDF, or other document that actually contains the text. Write from your knowledge of the text so that you do not replicate its language or syntax and instead produce a truly original presentation of the text's ideas. As you compose, you can check in with the original as needed to make sure you are representing the ideas accurately and did not unwittingly copy the text.

Remember that paraphrased text should always be cited. Even though you are not borrowing language when you paraphrase properly, you are borrowing ideas. This should be clear to your reader.

Example of Patchwriting: Since patchwriting may be a new concept to you, it will be helpful for you to examine an example of patchwriting as compared to an original paraphrase. We have adapted the following example from one given Angelika H. Hofmann's book Scientific Writing and Communication.

Here is the original passage containing ideas that a student writer wants to present to their reader. It is by an author with the last name "Smith."

> It has currently been recognized that both the type and characteristics of the rust layers formed on the steel surfaces are very important because they can determine their protective properties. According to a recent theoretical model developed by Chang et al., the long-term corrosion behavior of iron exposed to wet-dry cycles is largely controlled by the characteristics of the rust layers. Additionally, the differences between the corrosion behavior observed for different types of steels have been related to the rust layer characteristics.

Below is an example of patchwriting. Notice how if you compare each individual sentence below to Smith's original text above, they are only different in superficial ways, with small changes in word choice, word order, or syntax.

> It has been shown that the type and nature of the rust layers on the steel surfaces are very important because they determine their protective properties (Smith 2010). Recently, Chang et al. developed a theoretical model that shows how the long-term corrosive behavior of iron subjected to wet-dry cycles is primarily controlled by the nature of rust layers. The differences between the corrosion behavior seen in many types of steels have also been connected to the rust layer characteristics (Smith 2010).

Finally, below is an example of a true paraphrase. If you compare this paraphrase to Smith's original passage, you'll see author of this paraphrase clearly communicates and attributes Smith's ideas, but her presentation of them is fully original. This author has clearly understood what Smith has written and represents to her reader in her own words.

> Smith (2010) aims to study corrosive behavior in steel, including the way that steel type relates to corrosion behavior. Smith builds on previous research that has indicated a relationship between rust layer characteristics and corrosion behavior. More specifically, Smith takes up the model of Chang et al. (2008) to provide a theoretical approach for her research on how rust layer characteristics affect corrosion behavior.

Reference:
Hofmann, A. (2010). Scientific Writing and Communication: Papers, Proposals and Presentations (3rd ed.). Oxford University Press.

Plagiarism Type 4: Self-Plagiarism (also called text recycling)

Definition: At first, the idea of "self-plagiarism" may seem counterintuitive. If plagiarism is about using the work of other writers unfairly, then how could self-plagiarism exist? Self-plagiarism, sometimes called text recycling, happens when a writer copies and pastes phrases, sentences, paragraphs, or sections from one of their own previous works into a new work without clearly citing the previous work.

For professional academic writers, like your professors, there are a number of expectations for when they can and cannot reuse their own previous writing. For example, it may be acceptable to copy and paste the same sentences from an old grant proposal into a new one. However, it may not be acceptable to copy and paste the same sentences from one's previous book into one's new book without quotation marks and a citation of the earlier book.

For student writers, the goal of most writing is to learn. Every time we write something new, we have a new opportunity to learn both about the topic we are studying and about the practice of writing itself. For that reason, most professors disallow students from recycling writing they have done in one class into the writing they are doing in a new course. If you only wrote one paper in college and kept submitting it in every course you took, you would not learn very much.

In writing for courses, do not reuse your own previous writing without explicit approval from a professor. In writing for research purposes, consult with mentors about when and how it is okay to recycle your own words from other contexts.

How to Avoid: Avoiding self-plagiarism is relatively straightforward for student writers.

In coursework, if you are given a writing assignment and believe you have a compelling reason to reuse part of a text you authored in another context, you can discuss this possibility with your professor. It may be that the professor will approve some specific kinds of text recycling if they feel the new project you are working on still aligns with the learning goals of the current course. Otherwise, strive to explore new topics and questions with each paper you write.

In research, you may write several papers that focus on the same issue or problem. If you think text recycling might be a useful writing strategy, check with a mentor in the field about when and why text recycling is permitted in your field. In the absence of clear guidance from a mentor, avoid text recycling.

Plagiarism Type 5: Inadvertent Appropriation

Definition: Most academic writers fully intend to clearly attribute their sources, but even when a writer has good intentions, plagiarism can still occur. One example is inadvertent appropriation. This kind of plagiarism happens when a writer forgets that some of the ideas, processes, data, or words they are using originally came from another writer. This forgetting occurs when a writer abdicates their responsibility for keeping up with the sources of the ideas and language in their written work.

Inadvertent appropriation can happen in various ways. A writer might read another text without taking any notes on it and then mistakenly believe that ideas from that text are their own. Or, a writer might take notes on a text but do a poor job of distinguishing between which notes are direct quotes from a text, which are paraphrase or summary, and which are the notetaker's original ideas. Later, this writer might include another writer's words in their work without remembering that these words came from someone else.

At first, it might seem strange to think that you could get in trouble as a writer for appropriating other writers' work when you don't even know you are doing it. But academic readers believe that responsible reading practices are part of an ethical writing practice. You should expect to be held accountable for both how you read and how you write, because these practices are intertwined.

How to Avoid: To avoid inadvertent appropriation, the first step is to understand that the human memory is fallible. The tendency of people to read or hear an idea in one context and to later believe that they themselves generated the idea is so common that psychologists have a name for it: cryptomnesia.

Our memories become even less reliable when we are tired, under stress, or managing multiple tasks. You will have to confront all of these situations as a college writer.

The second step is to realize that a decision not to take organized notes on reading or to make best guesses about citation is a kind of sloppiness and lack of care that has ethical implications and serious consequences. When a student makes a weak effort at answering a problem in a problem set, the result will usually be a poor grade and a lack of learning. But when it comes to attributing work to its sources in your writing, a poor effort takes unfair advantage and violates the honor code.

The key strategy for avoiding inadvertent appropriation is to take careful, accurate notes when you do reading for courses and research. You must then consult those notes closely when you write. Develop a systematic approach to notetaking that works for you. You might use multiple columns

on a page to separate other authors' words and ideas from your own. Alternately, you might use text of different colors. Beyond such strategies, you should also record the author's name and page number as you take notes, so you can find your way back to the original text as needed and make necessary in-text citations. Regardless of how you do it, write your notes in a way that minimizes your reliance on your ability to simply remember which words and ideas belong to you and which have other sources.

Also, plan to read the texts you are writing about or borrowing from more than once. Academic readers often re-read a text many times during the writing process. This helps ensure they understand those texts, and it also makes it less likely they will inadvertently appropriate another writer's work.

Finally, cite as you write. Don't put citation off to be a final step in writing.

Type 6: Citation Errors

Definition: Citation errors are another type of plagiarism that arises due to negligence or lack of effort rather than an active intention to deceive a reader. Academic citation systems have very specific rules for how a writer shows the reader the source of a text. Sometimes writers lack knowledge or understanding of these rules and fail to seek out the guidance they need in order to follow the rules. In other cases, writers fail to do the careful proofreading needed to catch citation errors.

Citation errors can also occur when a writer gives the reader the wrong information about a text, making it difficult or impossible for a reader to understand which elements are borrowed and how to locate the original source.

One of the most common and problematic citation errors made by college writers is **under-citation**. For example, a student writer might be writing a whole paragraph paraphrasing ideas from an author named Henry Chang. Then, at the very end of the paragraph, the student writer cites Chang, assuming that the final citation tells the reader that all the above material came from Chang. Citation systems vary, but generally speaking, this approach fails to give the reader the information they need to understand the sources of all of the ideas, processes, data, and words in a text. The writer must direct the reader in **each sentence** in the paragraph as to whether it contains ideas from Chang, the student author, or both.

How to Avoid: Learning to correctly use the citation style systems that your readers expect you to use is a big responsibility. Different systems work in different ways, and there are a lot of details that writers need to understand how to use these systems correctly. Often, a professor will expect you to use a citation style without teaching you how it works. What should new college writers do to make sure they understand the system they are using?

The most popular citation style systems used by undergraduate writers are designed by scholarly organizations that publish books describing those systems in detail. These include the American Psychological Association (APA) style, Chicago style, and the Modern Language Association (MLA) style. These books contain detailed explanations of how these style systems work as well as models and examples of how to use them. These books are available in the Caltech Library and our Hixon Writing Center. They are invaluable resources for writers learning to use a style system.

Using these books can help answer many questions, such as how to cite a text you are working with in back-to-back sentences. In other cases, working with these books may help you pinpoint a particular question you are unable to answer. You can then take that question to your professor for guidance.

Citation management systems, like Zotero, which is endorsed by Caltech's Library, help writers keep up with sources and generate citations in a document. These systems can be very valuable time-saving tools for academic writers. If you use such tools to help you generate citations, be aware that they are only as good as the information the user gives them, and many of them still make mistakes. Proofreading for citation errors is a necessary practice for all academic writers.

Plagiarism Type 7: Common Knowledge Problems

Definition: One common way students plagiarize is to seek out a reference source, like Wikipedia, to answer basic questions about a topic. They then include that information in their writing without citation, because they believe it is common knowledge and doesn't merit citation. The question of what can be left uncited as common knowledge is one of the trickiest questions student writers face.

Let's think about an example. Most writers who mention that the Earth is a sphere would not cite a source for what shape the Earth is. The fact that the Earth is a sphere is an ancient discovery that is now commonly known. However, the matter might be handled differently in different academic writing contexts. A historian of science who is writing a paper about how and when the shape of the Earth was discovered would be likely to carefully document the sources of statements about how people understood the shape of the Earth over time. Similarly, an Earth scientist studying contemporary methods for measuring the exact shape of the Earth would also cite sources that have made specific claims about the shape of the Earth.

What do these examples suggest? Concepts that seem like common knowledge to outsiders may be objects of active discussion to experts in a field. Nearly anyone can tell you the Earth is a sphere—a clear case of common knowledge, right? Ask an Earth scientist if Earth is a sphere, and they will tell you that it's not, at least not exactly. For these experts, the shape of the Earth is an interesting and complex question that continues to be actively discussed.

There is another reason student writers should plan to cite common knowledge relevant to their writing topic. Students are novices in the fields they are studying and are encountering the principles of a field for the first time. In this way, what might be common knowledge to an expert is not at all common knowledge to a student. Attributing your sources is a way of telling your professor how you know what you know.

As a student, you should plan to cite ideas that you learn about from reference sources, particularly common knowledge relevant to your project. It is also advisable to consult with your professor about their specific expectations for how you cite common knowledge as well as material learned from the course lectures or lecture notes.

How to Avoid: If you only know something because you read about it in course readings or research for a course paper, plan to cite that knowledge. Some course guidelines for writing may offer you additional guidance. For example, one professor might indicate that they do not expect you to cite the textbook and that it is considered "common knowledge" for the course. Another

professor may prefer that you cite all ideas, data, processes, or language you borrow from a textbook. Pay attention to these guidelines and ask questions when ambiguous situations arise.

Another important practice is to avoid reading non-scholarly sources that you would not be willing to cite. Once you learn something from Wikipedia that you want to use in your paper, you become ethically bound to cite Wikipedia.

The penalties for overcitation in a course paper are typically minimal or non-existent. The penalties for omitting citations can be quite severe.

If you are writing a research paper for publication, you should consult closely with a mentor about the common knowledge question. What is considered common knowledge that does not require citation will depend on your field, the topic of your paper, the genre of your paper, and your intended venue of publication.

What about generative AI?

So far, this module has addressed the use and potential misuse of sources authored by humans. Until recently, all the texts that an academic writer might work with were authored by humans. Now, writers have access to generative AI (GenAI) tools, like ChatGPT and Google Gemini, that can be prompted to produce texts that lack a human author. These tools can also be utilized to revise or edit a text created by a student writer, which raises issues similar to those addressed above in the section on outsourcing and overcollaboration.

As a Caltech student, how can you use generative AI when doing coursework? **Students should follow course-specific guidance from each instructor about what uses of generative AI are allowed and which are not in that particular course**. You should expect your professors to have different expectations about the use of generative AI because their courses have different learning goals. One course might allow unlimited use of the tools, and another might prohibit any use at all. Another course could allow the use of the tools to write code, but disallow their use when writing an essay assignment. Under the Caltech honor code, you must abide by these policies in completing work for each course.

You will be expected to carefully review and adhere to course policies on the use of GenAI. If a course lacks a policy, it's a good idea to ask the professor what their expectations are for the ethical use of GenAI. If you lack guidance and no longer have time to ask the professor, assume any use of GenAI that is not explicitly required or encouraged in an assignment is not permitted. Be aware that GenAI tools are trained on huge amounts of text and may offer outputs that contain unattributed or incorrectly attributed ideas, processes, data, or language. LLMs are not reliable assessors of their own outputs, so when you use these tools, you, as the human author, must do your own due diligence to ensure you are not committing plagiarism if you utilize their outputs.

Questions? Email writing@caltech.edu to get support with any aspect of academic writing from the Hixon Writing Center.