Med Health Care and Philos (2015) 18:91–101
DOI 10.1007/s11019-014-9583-8

SCIENTIFIC CONTRIBUTION

Plagiarism in research

Gert Helgesson · Stefan Eriksson

Published online: 4 July 2014
© Springer Science+Business Media Dordrecht 2014

Abstract Plagiarism is a major problem for research. There are, however, divergent views on how to define plagiarism and on what makes plagiarism reprehensible. In this paper we explicate the concept of "plagiarism" and discuss plagiarism normatively in relation to research. We suggest that plagiarism should be understood as "someone using someone else's intellectual product (such as texts, ideas, or results), thereby implying that it is their own" and argue that this is an adequate and fruitful definition. We discuss a number of circumstances that make plagiarism more or less grave and the plagiariser more or less blameworthy. As a result of our normative analysis, we suggest that what makes plagiarism reprehensible as such is that it distorts scientific credit. In addition, intentional plagiarism involves dishonesty. There are, furthermore, a number of potentially negative consequences of plagiarism.

Keywords Fabrication · Intellectual contribution · Plagiarism · Scientific misconduct · Software · Scientific credit

Introduction

Plagiarism is a well-known and growing issue in the academic world. It is estimated to make up a substantial part of the total number of serious deviations from good research practice (Titus et al. 2008; Vitse and Poland 2012). For some journals it is indeed a serious problem, with up to a third of the published papers containing plagiarism (Zhang 2010; Baždaric et al. 2012; Butler 2010). Given that plagiarism is perceived as a considerable problem for the research community, spelling out in some detail what is to count as plagiarism becomes a matter of pressing concern. The technical development of software for detecting plagiarism also raises questions: What degree of overlapping constitutes plagiarism, and is overlapping all that matters?

Clarifying what constitutes plagiarism is one thing, and making clear what is wrong with it is another, although the two are interrelated. Are all forms of plagiarism equally bad? Are there perhaps even legitimate ways to plagiarise? If so, what makes plagiarism wrong?

In this paper we will mainly do two things. First, we will explicate the concept of "plagiarism", i.e. present an analysis of the concept aimed at further clarifying it. This means that we will look at previous uses of the term and through critical analysis come up with what we take to be an improved definition. While many organizations and research ethical guidelines present their definitions of "plagiarism", little work has so far been done in explaining and justifying the chosen definitions. Here we hope to make an important contribution. The point of the definition that we present is not to identify the essence or 'real nature' of plagiarism (we doubt that there is such a thing), but rather to extract one that is useful for the purpose of clarifying normative issues related to plagiarism, while being true to common uses of the term. Second, we will discuss plagiarism normatively, by taking a closer look at different aspects of it. We restrict our analysis to the context of research, since plagiarism in the arts, for instance, raise a partly different set of issues, and include partly different normative intuitions, which would require a separate analysis.

G. Helgesson (✉)
Department of Learning, Informatics, Management and Ethics, Stockholm Centre for Healthcare Ethics, Karolinska Institutet, 171 77 Stockholm, Sweden
e-mail: gert.helgesson@ki.se

S. Eriksson
Centre for Research Ethics and Bioethics, Uppsala University, Uppsala, Sweden

In order to evaluate an explication of "plagiarism" in relation to the present purpose, we first need to identify a set of conditions for adequacy. Although we will not systematically test suggested definitions against these conditions, they show what requirements our definition is intended, and believed, to meet to a reasonable extent.

Conditions of adequacy

The conditions of adequacy should identify relevant restrictions on any suggested definition for the definition to be reasonably adequate for the intended purpose in the intended context. Partly different criteria may become relevant depending on the intended use of the definition. We suggest, inspired by Brülde and Tengland (2003), that the following criteria for adequacy are relevant to a definition of "plagiarism" for our intended use:

- *Fitting language use*: The definition should not deviate too much from established language use, which is to say that it should catch basic semantic intuitions and should be able to handle paradigmatic cases—if acts that are usually considered to be instances of plagiarism are rarely taken to be so by your definition, then it fulfils this criterion poorly. The greater the number of such cases it covers, the better. However, it goes without saying that if there is no uniform language use, a logically consistent definition cannot cover all uses.
- *Precision*: The greater the precision of the definition, the better it is. Ideally, for each case the definition should settle whether or not it is a case of plagiarism.
- *Reliability (intersubjectivity)*: The definition is reliable if different users of it pass the same judgment on specific cases ("If plagiarism is defined as so-and-so, then this is (or is not) a case of plagiarism"). If a definition is reliable, then it produces the same outcome regardless of who is using it, which means that there is intersubjectivity in the use.
- *Theoretical fruitfulness*: The definition is more theoretically fruitful if it is better at distinguishing things that may be important to keep apart; it is better the greater the "job" it can do. For example, it is more theoretically fruitful if it can help to explain claims about plagiarism, such as why some instances count as plagiarism (or why some aspects are relevant for settling the issue) whilst others do not.
- *Relevance for normative purposes*: The definition should as far as possible identify as plagiarism those events that one would like to single out as morally problematic in this regard.
- *Simplicity*: The general idea that it is preferable for a definition to be homogeneous and ad hoc-free.

What is plagiarism?

Since it is important to determine what constitutes misconduct in scientific writing, and "plagiarism" is a much used concept in discussions of scientific misconduct, one could perhaps expect agreement and a fairly high level of precision regarding what constitutes plagiarism. However, while there is agreement about paradigmatic cases of plagiarism, there is less agreement regarding how plagiarism should be defined. In fact, the issue is rarely discussed in detail.

When the concept is explained in a recent newsletter from the US Office of Research Integrity, it looks deceptively simple: "It involves stealing someone else's work and lying about it afterward" (Sox 2012). Others prefer to speak of "copying" part of someone else's published work and using it without showing that it is borrowed from someone else. In the Longman Contemporary English Advanced Learner's Dictionary, the act of plagiarism is defined as "when someone uses another person's words, ideas, or work and pretends they are their own".

In the scholarly definitions, the more technical notions of "appropriation" and "credit" are central: "Plagiarism is the appropriation of other people's material without giving proper credit" (The European Code of Conduct for Research Integrity); "Plagiarism is the appropriation of another person's ideas, processes, results, or words without giving appropriate credit" (US Federal Policy on Research Misconduct). So the basic ideas seem to be that someone deliberately takes someone else's work, whether in the form of an idea, a method, data, results, or text, and presents it as their own instead of giving credit to the person whose ideas, results, or words it is. This is mirrored in the definition given by Merriam-Webster: "to steal and pass off (the ideas or words of another) as one's own: use (another's production) without crediting the source".

Two components of plagiarism

Common to these definitions is that plagiarism is composed of two parts: (1) to appropriate the work of someone else and (2) passing it off as one's own by not giving proper credit.

Let us first ask what it means to appropriate someone else's work. In some definitions, plagiarism is characterised as stealing. However, if plagiarism by definition concerns stealing, then it is not theft in the traditional sense of taking a thing, where if person A takes it from person B, then B will no longer have it. What is appropriated in such instances of plagiarism is *intellectual* property, as when people download copyright-protected films or music from the Internet. Thus, to the extent that plagiarism is theft, it is stealing someone else's intellectual work by copying.

Related to research papers, it is about copying another's text, tables, graphs, or pictures into one's own paper without having permission to do so (and with certain pretence, a point we shall be returning to presently).

We are, however, disinclined to include stealing in our definition. Although one may steal intellectual as well as non-intellectual property, and even talk about "theft of the recognition due to the original contributor" (Rathod 2012), talking about plagiarism as stealing is nevertheless misguided, at least as part of a definition. This is so because using someone else's text, say, and passing it off as one's own can be done regardless of whether one steals the text or not. One can do it by finding the text in a journal or book or by using an unpublished paper—or by stealing it from someone's computer or drawer. Thus, it seems that stealing is not a constituent part of plagiarising. In fact, you plagiarise a text even if it was willingly handed to you by a research acquaintance—if your use of it implies that it was you who created it. However, plagiarism does not preclude that the text presented as one's own has been literally stolen from someone else; you may steal a manuscript in order to plagiarise it (just as you may steal it in order to keep it without showing it to anyone). If you do, that means that you pass off the stolen manuscript as your own.

It may still be argued that there is a sense of "stealing" that concerns appropriating someone else's intellectual work and passing it off as one's own. In this sense you may steal someone's song if you play it and claim to have composed it yourself. This seems to mean that there is a sense of "stealing" that is equivalent to "plagiarising". If so, this second understanding of "stealing", which is distinct from the one discussed above, cannot contribute anything to a definition of "plagiarism". The conclusion remains: stealing, or theft, cannot be used as part of the definition of "plagiarism".

"To appropriate" does not have to imply stealing. It could also mean, for instance, acquire, borrow, take, or expropriate. We nevertheless suggest that "appropriate" should be avoided, just because it is such an ambiguous term and therefore would introduce obscurity in the definition. We instead suggest that "use" is employed.

It seems, then, that it is the second part of the definition that will distinguish cases of plagiarism from acceptable cases of using the results of another's intellectual effort. The second alleged aspect or component of plagiarism is passing it off as one's own. This can be done with or without the approval of the person or persons being plagiarised, so it is not about whether or not the re-use has the author's approval, but about what impression is given by that use. Using someone else's work and being dishonest or otherwise misleading about where it comes from seems to be what makes the act an act of plagiarism. But dishonest or misleading in a special way: If person A uses a passage from a text by B but claims that it was written by C, then, even though it is an incorrect claim, it is not plagiarism, but simply incorrect referring (if intentional, it is a kind of fraudulent behaviour). It is when A claims (explicitly or implicitly) to have written the passage him- or herself that it becomes plagiarism. This was brought out in the definition provided by Merriam-Webster above: it is when we *pass something off as our own*, although it isn't, that we plagiarise. This seems to be the core of plagiarism.

An intellectual product of one's own

It is no accident that plagiarism is discussed in relation to research, although it is also clearly relevant in relation to music, literature, art, and design, since it relates to using the product of someone else's *intellectual work* while passing it off as one's own. Note that there is no reason to restrict the use of the term to published work, since you may use someone else's work while passing it off as your own even if it is not published. For instance, you may do it by first stealing the manuscript from the author, by using passages from an unpublished manuscript circulated at a seminar, or by using ideas communicated at a lecture.

What if a person does not go to the trouble of writing up a paper in which the results of others' intellectual efforts are used with the pretence of being the person's own; what if the person simply makes the wrongful claim that "this is my work"? Would that also be plagiarism? Example: A scientific paper in astrophysics is published in a renowned journal by a group of researchers. Researcher Ynotme, not part of the group, then goes public falsely claiming that the published results are hers. Would she thereby be plagiarising? Our explication so far leans towards the view that plagiarism concerns *a product of one's own*, containing the appropriation of the intellectual work of someone else. We believe that it would be constructive to claim that plagiarism consists not only in passing someone else's work or intellectual product off as one's own, but also in using it *as* a product of one's own. Going for this position, falsely claiming a work of another's to be one's own would not be plagiarism, but would count as a false accusation of plagiarism and theft.

Our definition of "to plagiarise" would, thus, at this stage be: to *use* someone else's intellectual product while passing it off as one's own, where "use" is meant to indicate that it is made part, or the whole, of a product of one's own. However, although quite a few attempts at a definition of "plagiarism" include elements such as lying or pretending it is one's own intellectual work, others rather describe the second part of the definition in terms of not giving the proper or appropriate credit. While the first set of expressions—lying, stealing, and pretending—implies *intention*, the second set is neutral in this regard.

While ordinary language use of "plagiarism" certainly allows for the act of plagiarising being intentional, it seems as clear that plagiarism does not necessarily involve an intention to deceive. We therefore would like to suggest a definition that does not require intention. The notion of "passing something off" also seems to imply intention, and therefore ought preferably to be avoided. A slightly modified definition, where we define "plagiarism" rather than "to plagiarise", would therefore read: *Plagiarism = an instance of someone using someone else's work, thereby implying that it is their own.*

Plagiarising ideas? Plagiarising work?

What, then, counts as an intellectual product? The standard case of plagiarism is the use of someone else's text. We have seen that Merriam-Webster mentions "words or ideas", while Longman talks of "words, ideas, or work". Is it reasonable to say that ideas can be plagiarised—and what about work? Let us look at ideas first.

It seems that one can talk about plagiarising ideas just as well as one can talk about plagiarising research results or text, since ideas are obvious examples of results of intellectual work. If someone uses another's idea and implies that it is an idea of their own, that someone is plagiarising. True, it must be admitted that it may often be much more difficult to verify that an idea has been plagiarised compared to research results or text. Ideas are not always documented, but might be presented at conferences or in personal conversation, etc. The difficulty pertains both to finding out about the plagiarism and to making a convincing case for idea plagiarism to have taken place. There is no clever software to discern this, nor is it easily proven that an idea is not independently arrived at. These difficulties are, however, practical; they do not change the fact that ideas can be plagiarised.

Some might be reluctant on ideological grounds to accept that ideas can be plagiarised. They might think that ideas should be free and not be the intellectual property of anyone. However, our position is agnostic on this ideological debate since the definition does not rely on notions of stealing intellectual property. Freedom of ideas is compatible with the view that you are plagiarising if you use someone else's idea while implying that it is your own.

What, then, about plagiarising work? As previously noted, plagiarism in relation to work must concern intellectual work. In this context, the term "work" has two distinct senses: a product based on intellectual labour or that labour itself. When someone is plagiarising a text presenting research results, thereby implying that they are presenting their own results, then that person also implies that they have done the work leading up to the results. In that sense you can say that the person is also plagiarising the work put into it. By plagiarising someone's idea, you, by the same token, make implicit claims about the work leading up to that idea.

But it is hard to see that it makes sense to talk about plagiarising work (labour) directly. Let us look at an example: Say that Mr A visits Ms B and sees a beautiful chair that Ms B has made to her own design. Mr A goes home, builds an identical chair, and claims when friends ask that it is of his own design. When it comes to the chair, it is clear that it is the idea of making the chair just like that, i.e. the design, and not the work of making the chair (which he in fact did), that is plagiarised. Plagiarising work means plagiarising ideas relating to how to do the work, the results of work, or the documentation of how the work was performed, not the labour itself—the latter would be to *repeat* the work, not to plagiarise it.[1] We therefore choose not to talk about work, but instead of an intellectual product being plagiarised. So, our definition will be the following.

> Plagiarism = *def.* an instance of someone using someone else's intellectual product (such as texts, ideas, or results), thereby implying that it is their own.

Demarcations: self-plagiarism et cetera

Plagiarism being part of the standard definition of research misconduct, and therefore often regulated, allegations of plagiarism are more likely to be investigated than many other potential instances of deviations from good research practice. If it can be shown that other problematic behaviours can be covered by the definition of plagiarism, this will help make them eligible for investigation. Anekwe has in this way suggested that honorary authorship and ghost-writing[2] are instances of plagiarism, since these practices entail claiming merit for work done by others, even if those others condone the act (Anekwe 2010). It follows from our definition that we can agree with his conclusion.

It has become increasingly common to discuss so-called self-plagiarism as a special case of plagiarism (Roig 2006; Brogan 1992; Samuelson 1994). Perhaps this is prompted by a similar wish to include such behaviour in what can be

[1] Only if the result of intellectual work is a novel idea about a way to process a certain task (a method) will it be possible to plagiarise by *repeating the processes* and not disclosing where the idea of doing it like *that* originated. Which is to say that (the idea of) a method may be plagiarised by using it and not disclosing that someone else came up with it, thereby implying that you invented it yourself.

[2] It is, of course, not the writing that constitutes plagiarism in the context of ghost-writing, but the claim to have written or co-authored a text completely written by others.

reported and investigated. However, there is a considerable difference between plagiarism and self-plagiarism, in that plagiarism involves presenting the results of someone else's intellectual efforts as one's own (as is reflected in the different definitions discussed above), while self-plagiarism does not involve the work of others but is restricted to the reuse of one's own material. Similarly, if considered in the light of our explication of plagiarism, self-plagiarism clearly cannot be described as "using someone else's intellectual product, thereby implying that it is one's own". Therefore one might argue that self-plagiarism is a contradiction in terms, and thus a confusing way of raising the problem of redundantly overlapping publication (Bruton 2014 with many references).

Still, there are similarities between plagiarism and cases described as self-plagiarism. Both usually involve redundant publication—in both cases, new efforts and findings are quite often implied when in fact previous research has just been republished, with the consequence that scientific credit is obtained twice (or more) for something that deserves credit once only. Perhaps this is the greatest perceived similarity: in both plagiarism and so-called self-plagiarism, researchers are acquiring undeserved credit for research. Inspired by this, one might be inclined to suggest a definition that differs from the ones discussed above, stressing the "undeserved credit" aspect, such as: Plagiarism = *def.* an instance of someone's acquiring undeserved scientific credit, either by using someone else's intellectual product, thereby implying that it is one's own, or by presenting one's own previously recognized work as new.

However, this definition has some important weaknesses compared to the one we propose. First, it deviates from the vast majority of definitions of plagiarism, since it doesn't require that *someone else's* intellectual product is involved. Thus, it fits established language use poorly. Furthermore, it makes plagiarism hinge on whether or not undeserved scientific credit was in fact acquired, which is irrelevant in ordinary language use—it is still plagiarism, even if the submitted paper containing unacknowledged copied material does not get published. One might try to counter this weakness by adding "or trying to acquire" after "someone's acquiring". But that trick does not work; because it may be an act of plagiarism even though the plagiarizer does not succeed in acquiring undeserved credit, nor tries to do so (he may not know that the cut-and-paste method is unacceptable). Neither *acquiring undeserved scientific credit* nor *trying to do so* is a necessary component of plagiarism. The basic error in this attempt at a definition is that it puts focus on the wrong thing, namely on the *effect*, or the intended effect, of doing something rather than on the very act that the definition concerns. This will no doubt have implications for the theoretical fruitfulness of the definition, as well as for its relevance for normative purposes. Furthermore, although of lesser importance, by containing two distinct components, this definition is not as simple as the one we propose. This lack of simplicity means that in many situations it will be unclear what happened when we learn that "P plagiarised", since it may be either that P used someone else's intellectual product or reused his/her own. For clarity, it is preferable, all else equal, that definitions do not have the form "A is defined as this *or* that".

In many research areas papers are co-written. If one of the authors reuses text without proper notification, thereby implying that what is written is his/her own, then this is primarily a case of plagiarism, not self-plagiarism, because here we have an individual claiming by implication to be the sole author of that which is the result of an intellectual effort also made by others.

It can be helpful to distinguish plagiarism from duplicate publication, text recycling, salami slicing, and copyright infringement (Bruton 2014; Roig 2006). While we define plagiarism as "using someone else's intellectual product, thereby implying that it is one's own", self-plagiarism is sometimes better described as duplicate publication. Duplicate publication concerns publication of whole articles or texts (or sets of data or results) more than once without proper notification of this fact. When the "self-plagiariser" uses shorter passages of texts (or some figures, etc.) in repeated instances, we prefer to speak of inappropriate recycling of material. When the same study or set of experiments is dispensed in small chunks in different papers just to increase the number of publications, we have what is commonly known as "salami slicing".

Plagiarising someone else's intellectual product is not the same thing as infringing on someone's copyright. This follows clearly from our definition. The results of intellectual endeavour can be plagiarised without intellectual property claims being involved; for instance, it is perfectly possible to pass off as one's own a text of unknown origin from the dim and distant past. It is also possible to infringe someone's copyright without plagiarising. To publish an illustration owned by others or a passage of text that contains a large number of words might, proper referencing notwithstanding, be an infringement if in fact you need the owner's permission to publish. A further difference is that ideas can only be protected by copyright if given a tangible form (if they are written out) while they can be plagiarised even if only communicated orally. Yet another is that copyright protects the economic interests of the copyright holder while a do-not-plagiarise principle protects due recognition. To sum up, these instances of improper handling of material can co-exist in the same act and occur separately (they neither imply nor rule out each other).

Irrelevant aspects

It is sometimes asked whether certain aspects other than the ones discussed above are relevant in determining whether or not a certain act is an instance of plagiarism:

- the scientific merit to be gained from the publication
- the locus of the plagiarism (for instance, a published paper, a student essay, a summary of a doctoral thesis, or an oral presentation at a seminar)
- who is plagiarised
- the intended audience or purpose of the plagiarising work as compared to the original

One idea that we have encountered is that something that would be considered plagiarism if appearing in a published scientific paper may not be considered plagiarism if, for instance, appearing in a report ordered by a public authority or in a student paper not intended for publication. However, the locus of the plagiarising work or how conducive it is to career promotion is irrelevant to whether it is plagiarism, although that may be relevant to an evaluation of how serious the misconduct is; for instance, one may argue that the greater the research merit of a paper containing plagiarism, the more serious it might be considered, analogously to how theft may be considered more serious the more money that is stolen.

If something is to be considered plagiarism or not is also independent of who is plagiarised. For instance, it makes no difference if the person plagiarised is considered insignificant. It is also irrelevant to the evaluation of whether or not there is plagiarism if the plagiarised and plagiarising texts are used for different purposes, have different intended readers, or are of very different dimensions. It is still plagiarism if someone copies something from a short research paper and includes it in an extensive book. Whether or not the plagiarised text was published in an indexed, peer-reviewed journal is equally irrelevant.

Plagiarisers sometimes defend their actions by referring to cultural differences in the attitudes towards the work of others, and especially work of authorities (Sun 2012; Chandrasoma et al. 2004; DeVoss and Rosati 2002). They refer to an attitude that, out of respect, one must not meddle with the thoughts of great thinkers by re-writing their work—it should be left as it is. If such an attitude of respect, or even reverence, also exists in relation to research, this is at variance with the scientific ethos that is assumed all over the world: there should be no exemptions for local idiosyncrasies. However, using quotations to a reasonable degree is in accordance with good scientific practice as long as quotation marks or indentations with correct references are used.

It has happened that researchers with insufficient skills in the English language have been encouraged by their supervisors or colleagues to use another paper as a template and change data in order to include their own results instead of those in the template, with a considerable text overlap as a consequence (Couzin-Frankel and Grom 2009). Regardless of what the underlying explanations are, they have no bearing on whether or not a certain act is an act of plagiarism.

Does size matter? Or only originality?

Plagiarism can be more or less extensive, ranging from whole chapters of books, or entire academic papers, to shorter passages. Is there a lower limit to what counts as plagiarism? If so, when does it cease to be plagiarism—is it if it falls short of a certain number of copied words in a line or of a certain percentage of overlapping text in an essay, or does it depend on something else, such as the quality of that which is reused without notification?

We will argue that to the extent that quantity matters, it depends on whether quantity has an impact on quality. This is to say that quantity matters only indirectly, while quality matters directly (i.e. as such). Plagiarism may consist in very short passages of text. In principle, it may consist in one word or expression only. But that would have to be a very special, novel word or expression creatively used, e.g. for naming a new concept, perhaps something that throws new light on an area of interest. It would also have to be a situation where the plagiariser, by plagiarising, gives the impression that s/he invented the concept.

Using ordinary words like “and”, “it”, or “are” can never, as such, constitute plagiarism. Nor can the use of series of words, or sentences, which are so ordinary that they cannot meaningfully be ascribed to anyone. Examples: “He saw me”, “Open the door”, or “I am tired and need some sleep.” Due to their commonness, they belong to a common pool of expressions and sentences to which no one has an intellectual claim. If a number of people independently have “created” the same expressions, these expressions *ipso facto* lose their exclusivity. Such word combinations *cannot* be plagiarised (or so we would like to argue) because they cannot be considered to have been taken from someone else (in particular). This means that there is no case of plagiarism if such sentences have been copied and pasted from another text without (ordinarily due) notification. The Committee on Publication Ethics (COPE) gives the example of “smokers with chronic obstructive pulmonary disease” being used in science as a standard phrase, having more than 58,000 hits on Google (Wager 2011). Other examples are “The questionnaire was distributed to a random selection of…”, “Statistical analysis was conducted using SPSS …”, and “The study was granted ethical approval by the ethical review board in …”.

It is our firm belief that researchers can produce many more examples of this kind from their respective fields.

However, one reservation is called for. If a sufficient number of ordinary sentences not really belonging to anyone are put after each other in exactly the same way as by another author, then this may again be considered to be plagiarism. The longer passage may uniquely be attributed to a particular author, even though the individual sentences cannot.

Nor is it plagiarism to refer to others' results by stating numbers, like percentages, that express the results, without putting them between quotation marks. The same goes if someone, with references, states analysis categories identified in someone else's qualitative study. This is so because, in the case of the numbers and categories, adding quotation marks does not make things any clearer. If there is any reason to believe that some readers will hesitate about whether you named the qualitative categories yourself, while you intended to state them literally from the cited paper, then you might need to be more explicit about this or add quotation marks after all.

To summarise, we claim that plagiarism (in principle) can consist in as little as one word, while there are many standard sentences describing research methods that will not be plagiarism even if, in fact, copied from someone else. This is to say that the unmarked reuse of some very short passages might be plagiarism, even though the reuse of other equally short passages would not. The conclusion to draw from this is that plagiarism has to do with quality rather than quantity—or, more precisely, with what is unique rather than so common that it cannot be attributed to anyone.

Plagiarism and probabilities

When it comes to investigating accusations of plagiarism, failing direct proof, the investigation will have to rest on probabilities. The longer and the more unique the identical passage, the greater the likelihood of its having been plagiarised. Still, if fairly ordinary passages or sentences, which are not common enough to be considered as not belonging to anyone, are in fact copied from someone else without use of quotation marks, then they *are* plagiarised, even if, for lack of further evidence, they will be considered by an investigator as not plagiarised. The criterion for plagiarism does not involve probabilities. Probabilities become relevant as part of a decision method when trying to settle whether or not an act of plagiarism has been committed. Also, if exactly the same, non-trivial sentence is written independently by two different authors, then there is no plagiarism involved, even if it will seem unlikely to an investigator that it was not. It does not become plagiarism because it seems to be plagiarism.

It is important to notice that software used to identify plagiarism—like iThenticate, Viper, and Turnitin (Khan 2011)—only detects text similarity. Such software can certainly be of help in detecting potential cases of plagiarism, but does not, of itself, identify plagiarism. In most cases, a proportion of overlapping text, expressed in a certain percentage, is insufficient to settle whether or not plagiarism is present. If you have 100 % overlap, then you know. You can also strongly suspect plagiarism if you find an overlap exceeding, say, 70 % for the entire text. But using a certain percentage over an entire paper, as some scientific journal editors do, seems to be a shaky foundation for deciding whether or not to investigate plagiarism. For instance, for a four-page paper a completely copied half-page would render a 12.5 % rate for the entire paper. To copy a third of a page word for word in a four-page paper, which in most cases would suffice for a convincing case of plagiarism, would render an overall overlapping rate of only 8.33 %. Clearly, then, you cannot use an arbitrary cut-off point of say, 20 %, because that would potentially miss many an instance of plagiarism.

Furthermore, this software can only help to identify plagiarism of text or numbers, whereas it is useless if instead what is plagiarised is ideas. It is also sensitive to language, which means that it does not detect plagiarism resulting from, say, taking a text written in English and using it, translated, in a text in German or French.

The normative status of plagiarism

It is part and parcel of good research practice to know the difference between plagiarism and established rules for citations and quotations. But judging the normative status of different cases of plagiarism is another issue. While some will constitute major cases of misconduct, others may be considered minor deviations from good research practice. While copying half a research article into one's own paper would be serious misconduct, copying 5–10 average-length and spread-out sentences of limited importance in a five-page paper is perhaps not. Yet another issue (left aside in this paper) is what policies are reasonable to adopt at universities and research institutes.

Before discussing what makes some forms of plagiarism worse than others, we should say something about what makes plagiarism bad to begin with. Plagiarism is one of the "core" instances of research misconduct, the other two being fabrication and falsification. "Fabrication" concerns making up research results instead of actually producing them by doing research. "Falsification" concerns tampering with research results, research methods, or data analysis. Common to both is that the researchers are misleading about what they have accomplished—they

pretend to have done the research, to have reached the presented results, to have used the correct methods and procedures, or to have applied appropriate analyses in the way described. Fabrication and falsification are directly detrimental to science, with the indirect effect that researchers may prosper from publications based on these kinds of fraud.

Plagiarism is commonly held to be reprehensible because it makes publications (etc.) misleading regarding who deserves credit for the intellectual work done—thus, it is unjust. It is also common to refer to the very act itself, declaring it to be an instance of cheating and betraying, both reprehensible acts. Some also point to the person plagiarising, maintaining that an additional wrongness of plagiarism lies in the fact that it makes the person a cheat and an impostor. These remarks, however, are restricted to intentional plagiarism. Plagiarism may, further, have unjust consequences by affecting who gets good grades, academic positions, and research funding.

In addition, plagiarism of data or results distorts the scientific record by giving a misleading account of research accomplishments. What is presented as new collections of data or as new results is not—instead it is a reiteration of what has already been done. Thereby it also distorts meta-analyses.

Let us now ask what aspects affect the normative status of a case of plagiarism. There is no direct connection between what aspects are relevant to determine whether or not something *is* plagiarism and what aspects are relevant to a specific normative judgment of an instance of plagiarism. For instance, aspects that are relevant when determining whether or not something is plagiarism may have nothing to add when it comes to evaluating gravity, as might sometimes be the case when regarding the originality versus ordinariness of passages appropriated in material of one's own. Other aspects are irrelevant when determining whether or not it is a case of plagiarism, but may be relevant when determining the seriousness of the plagiarism, such as whether or not the plagiarism was intended or the scientific merit value of the publication or presentation containing plagiarism. So what aspects affect the normative status of a case of plagiarism? Candidates include:

- the value of that which is appropriated
- the manner in which the plagiarism is performed
- the degree of harm to the plagiarised person(s)
- the degree of personal gain to the offender(s)
- whether the plagiarism is intentional or not

Before going any further, we should first note that one may distinguish between what makes plagiarism worse *qua* plagiarism (roughly points 1 and 2), what makes an action involving plagiarism worse on the whole (points 3 and 4), and what makes the plagiariser more or less blameworthy (point 5).

Intentional or unintentional?

To begin with the last point, a case of plagiarism is judged differently depending on whether or not the offender did it on purpose, just as other intentional wrongdoings are considered more blameworthy than unintentional ones. Sloppy quotation practices, or ignorance, are not as blameworthy as intentional fraud. If someone falls prey to *cryptomnesia*, i.e. unconscious plagiarism that happens when you remember the idea but not that you got it from someone (Roig 2006), they might to a certain extent be excused. However, one might argue that the very plagiarism is equally bad whether intended or not, while it is more reprehensible to plagiarise intentionally than otherwise.

Even though a case of plagiarism is judged differently depending on whether the offender did it on purpose or not, one may also be held responsible for one's ignorance. Good research practice involves knowing where to draw the line between acceptable and unacceptable behaviours relating to research, such as the unacceptability of fabricating or tampering with data or results. That it is unacceptable to cut and paste from other people's published work, without stating the source and showing exactly what passages are quoted, is part of that required knowledge of good research practice. If one has not been taught this, part of the blame for plagiarism must fall on one's teachers and supervisors. Excuses nevertheless cannot eliminate the fact that disrespecting standard rules of quotation is a deviation from good practice. Someone who is plagiarising is always blameworthy, to the extent that readers thereby are misled about who deserves credit for the work.[3] Still, those who mislead intentionally are more blameworthy.

Wrong as such and bad consequences

Let us return to the previous points. While the first two points concern the disvalue of plagiarism as such, the other two points concern the consequences of plagiarism. A case of plagiarism might be considered graver if the material plagiarised had the potential for greatly benefitting the originator economically or by having great impact. For example, a plagiarising publication can rule out the possibility of obtaining a patent. Conversely, we would perhaps think it worse to build a whole well-renowned career upon plagiarising others, than to plagiarise in a way that never brought any particular advantages.

[3] It should be noted that it does not have to be the authors' fault that a paper is misleading about who deserves credit. Leonard Fleck has brought to our attention instances of journals, unbeknown to the authors, having mistakenly removed references or quotation marks in the text, causing the text to give the impression that some phrases quoted from others are the authors' own.

As can be seen in these examples, what harm is caused to the plagiarised person(s) and what the offenders gain in a specific case depends not only on the characteristics of the specific plagiarism as such but also on things that lie beyond that (such as the reception of the alleged work, or the legal consequences of making ideas public). Depending on these "external" circumstances, the harm or benefit may be limited or great. This supports the idea that it is meaningful to distinguish between judging the plagiarism *qua* plagiarism and judging the act of plagiarism and its consequences as a whole.

What, then, makes plagiarism *qua* plagiarism more or less reprehensible? First, it normally makes a moral difference what is plagiarised: is it an idea, or research data or results, or is it rather useful phrases or background descriptions? Plagiarism of research results, and also discussion, is seen by many as considerably worse than plagiarism from the introduction or methods sections. One rationale for this view is that plagiarism of data or results involves fabrication (the offender gives the impression of presenting new data/results while this in fact is not the case). This means that plagiarism of data or results is worse than simple textual plagiarism because it also involves something else that is bad (fabrication).

Second, plagiarism of data/results may be considered worse as plagiarism because it involves something more novel, more creative, and thus scientifically more valuable than background and methods sections normally do; plagiarism of the latter often only involves free-riding on the labour and phrasing skills of others. There may, of course, be exceptions to this rule: the background section may present previous research endeavours in a new and eye-opening way likely to revolutionise future research; and the methods sections of methodological papers do indeed tend to contain their most novel and creative work. Regardless of where the main merits of a paper are located, plagiarism of those parts is more reprehensible than plagiarism of less important parts. To sum up, the greater the value of that which is plagiarised, the graver the plagiarism. The value we have in mind stems from the novelty and potential of affecting knowledge development within the specific field. Arguably these are also the aspects that are most likely to affect the scientific credit to be gained from the publication (regardless of whether it is the original or the plagiarising piece).

Lastly, the very act of plagiarism might be perceived differently according to the manner in which it was performed. If someone has a reference to where the material has been lifted but neglects to use proper quotation marks, it might be a sign of not having the *intention* to deceive, and we therefore find the act less reprehensible. But the *very act* also seems less objectionable in this example, since at least some merit is given to the original source and because readers are able to check the source, which they otherwise wouldn't. How you do it thus plays a part in evaluating the seriousness of the offence of plagiarism.

Pre-comprehension and being misleading

Within some research fields in medicine and the natural sciences, it seems to be quite common when writing up a paper to state certain key passages by recognised authorities, particularly in the methods section, with references but without showing that it is a literal quotation. It also seems common to recycle text literally from methods sections of one's previous papers without quoting. This practice is sometimes taken so far as to use whole sections verbatim over and over again without proper citation practices being followed. For example, research groups may write up what they believe to be the perfect method section, and then it is a given that this section is used in any paper originating from the group. Those who are used to this argue that everyone knows about it and that the practice therefore is not misleading.[4] One might, of course, question this; for instance, when papers are attracting a wider audience, these readers cannot be expected to know of the particular authoritative text sections taken for granted by the insiders.

It would, of course, be easy to change the practice, if there is such, and abide by standard rules for quotations, for instance, by adding quotation marks to the quoted passages. But is it reprehensible to leave them out? What position one takes on this issue may have to do with academic background. In areas where the written word is central and where researchers are used to considering the entire paper as the results of the research, people will not be inclined to take quotation rules lightly, while there seems to be a partly different attitude within medicine and the natural sciences, where what is considered to be the research results is that which you find in the results section of the paper, primarily in the tables or expressed in mathematical form. The discussion section is then usually also considered to contain material that is clearly the researchers' own contribution, while little sentimentality or personal strings are felt regarding background and methods sections (unless it is a methodological paper).

One could point out that it would be a disservice to science to change the description of a commonly used method every time a publication is prepared, just to avoid charges of plagiarism, if doing so results in a less comprehensible text. So if everyone knows of this practice, more is gained by reproducing the methods section

[4] Our claims here regarding practices are based on anecdotic evidence only. However, based on our teaching about 500 doctoral students per year, and having heard this frequently in class, we believe this to be fairly common, or at least far from unique.

verbatim than what is lost. Still, why not do this *and* follow proper citation practices?

This example illustrates that whether certain behaviours are misleading or not partly depends on the pre-comprehension of those receiving the message. What you ought to do, then, is dependent upon what the pre-comprehension actually is, or what it reasonably can be expected to be.

We noted above that some non-native speakers of English defend their actions by reference to their using others' papers as templates, and we said that this response has no bearing on whether their acts are plagiarism or not. Might this line of reasoning regarding pre-comprehension be used in defence? In 2007, a letter to the editor appeared in Nature that defended this practice both by downplaying the importance of anything but the results and by reference to the commonness of this practice. The author of the letter also suggested that when borrowing sentences makes a non-essential section better, this should not be considered plagiarism in a normative sense (Yilmaz 2007), a statement which some scientists seem to agree with (Pecorari 2012).

There are several possible responses to this. One is that if someone else's text is used as a template without the fact being duly noted, then this will very likely constitute plagiarism. Unless there is an open agreement beforehand that certain texts are free to use as templates, the practice is reprehensible. Another response is that if scientists do not have a working skill in English, then it would be better if they wrote their papers in their native language and had them translated by professional translators.

There are some other important things to note as well, to which we now turn.

Copyright and the risk of getting reported

Even if the use of certain key passages by recognised authorities without following established general referencing practices is recognised as good research practice in a certain context, there are two circumstances that argue strongly for a cautious approach to such a practice. They both point out that the suggested practice only "works" as long as you do not get exposed to the wider research society practising it. First, to reuse, for instance, a widely known methods section might not fool anyone in the field about its origin, but it might still be wronging the publishing houses involved. The copyright in the original text is likely to be owned by someone, and if someone else uses the text without proper referencing then that person will be infringing the copyright. Also, the publisher of the text will expect all material to be original unless the contrary is explicitly stated or shown. If the author or authors are not open about this, the publisher will be deceived. Second, to have sections imported from other sources without proper references and quotations is to invite accusations of research misconduct from those who spot the practice and are willing to cause harm to those doing it. We thus advise against this practice on these grounds.

Conclusions

We suggest that plagiarism should be understood as "using someone else's intellectual product (such as texts, ideas, or results), thereby implying that it is their own". This may be done intentionally or unintentionally. This fits the use of the term in ordinary language fairly well, while at the same time being sufficiently precise. Arguably it is reliable by being simple and easily comprehensible. We suggest that our discussion supports the view that the definition is theoretically fruitful and highly relevant for normative purposes. As a result of our normative analysis, we suggest that what makes plagiarism reprehensible is that it involves an unfair acquisition of scientific credit. In addition, intentional plagiarism involves dishonesty. In plagiarism of data or results, fabrication is also implied.

Acknowledgments We would like to thank the participants at seminars at Stockholm Centre for Healthcare Ethics, Centre for Research Ethics and Bioethics at Uppsala University, and at the International Bioethics retreat in Paris 2013 for valuable suggestions and constructive criticism of earlier versions of this paper.

References

Anekwe, T.D. 2010. Profits and plagiarism: The case of medical ghostwriting. *Bioethics* 24(6): 267–272.

Baždarić, K., L. Bilić-Zulle, G. Brumini, and M. Petrovečki. 2012. Prevalence of plagiarism in recent submissions to the Croatian Medical Journal. *Science and Engineering Ethics* 18: 223–239.

Brogan, M. 1992. Recycling ideas. *College and Research Libraries* 52(5): 453–464.

Bruton, S.V. 2014. Self-plagiarism and textual recycling: Legitimate forms of research misconduct. *Accountability in Research: Policies and Quality Assurance* 21(3): 176–197.

Brülde, B., and P.-A. Tengland. 2003. *Hälsa och sjukdom: en begreppslig utredning (Health and disease: A conceptual inquiry)*. Lund: Studentlitteratur.

Butler, D. 2010. Journals step up plagiarism policing. *Nature* 466(7303): 167.

Chandrasoma, R., C. Thompson, and A. Pennycook. 2004. Beyond plagiarism: Transgressive and nontransgressive intertextuality. *Journal of Language, Identity and Education* 3(3): 171–193.

Couzin-Frankel, J., and J. Grom. 2009. Plagiarism sleuths. *Science* 324(5930): 1004–1007.

DeVoss, D., and A.C. Rosati. 2002. "It wasn't me, was it?" Plagiarism and the web. *Computers and Composition* 19: 191–203.

Khan, B.A. 2011. Plagiarism: An academic theft. *International Journal of Pharmaceutical Investigation* 1(4): 255.

Pecorari, D. 2012. Textual plagiarism: How should it be regarded? *Office of Research Integrity Newsletter* 20(3): 3,10.

Rathod, S.D. 2012. Plagiarism: the human solution. *Office of Research Integrity Newsletter* 20(3): 1,7.

Roig, M. 2006. Avoiding plagiarism, self-plagiarism, and other questionable writing practices: A guide to ethical writing. Office of Research Integrity 2006. www.cse.msu.edu/~alexliu/plagiarism.pdf.

Samuelson, P. 1994. Self-plagiarism or fair use. *Communications of the ACM* 37(8): 21–25.

Sox, H. C. 2012. Plagiarism in the digital age. *Office of Research Integrity Newsletter* 20(3): 1,6.

Sun, Y.C. 2012. Does text readability matter? A study of paraphrasing and plagiarism in English as a foreign language writing context. *The Asia-Pacific Education Researcher* 21(2): 296–306.

Titus, S.L., J.A. Wells, and L.J. Rhoades. 2008. Repairing research integrity. *Nature* 453(7198): 980–982.

Vitse, C.L., and G.A. Poland. 2012. Plagiarism, self-plagiarism, scientific misconduct and VACCINE: Protecting the science and the public. *Vaccine* 30(50): 7131–7133. doi:10.1016/j.vaccine.2012.08.053.

Wager, L. 2011. How should editors respond to plagiarism? COPE discussion paper. 26th April, 2011. http://publicationethics.org/files/Discussion%20document.pdf.

Yilmaz, I. 2007. Plagiarism? No, we're just borrowing better English. *Nature* 449(7163): 658.

Zhang, Y. 2010. Chinese journal finds 31% of submissions plagiarized. *Nature* 467(7312): 153.