

RESEARCH ETHICS

2020.05.18

JUHO KIM



CLASS ON MAY 20TH

- X No presentation workshop
- X Replaced with office hours for 1:1 chat with instructors
- X *“We can talk about your presentation 1:1, any other topic related to this class, life as a researcher in general, or if you feel comfortable, about life or anything you like!”*
- X If you want a slot, leave a comment on Campuswire by 11:59 tonight.



TAKE-HOME MESSAGES

- X Ethical issues apply to all phases of research.
 - Know what they are, learn to properly address them.

- X Ethical implications exist in all areas of computing research.
 - Know what they are, learn to properly address them.

PLAGIARISM & CREDIT



WHAT IS PLAGIARISM?

- X Wrongful appropriation and stealing and publication of another author's "language, thoughts, ideas, or expressions" and the representation of them as one's own original work. [Wikipedia]
- X Serious academic dishonesty
- X Self-plagiarism
 - Reusing your own published work in subsequent publications without citing the original source
 - Duplicate publication



meg price

@meggy97

Follow



i would pay good money for someone to write my research essay for me

11:57 PM - 3 Jan 2018

3 Retweets 30 Likes



28 3 30



Essay Writers @heroespapers3 · Jan 4

Replying to @meggy97

Get essay help fast and easily
Confidentiality guaranteed
No grammatical errors and spelling mistakes
No run-on sentences
Website: heroespapersinternational.org
Email: heroespapers@gmail.com

28 3 30



AppleWriters   @apple_writers · Jan 4

Replying to @meggy97

Get the perfect essay tailored to your academic needs. #Writing, #editing, and #proofreading services. Turnaround from 3 hours. Visit goo.gl/nwD9iD to place an order

28 3 30



Guruwriters @guruwriters · Jan 4

Replying to @meggy97

hmu



@superiorpapers8 @superiorpapers8 · Jan 4

Replying to @meggy97

Hit me up, superiorpapers334@gmail.com.

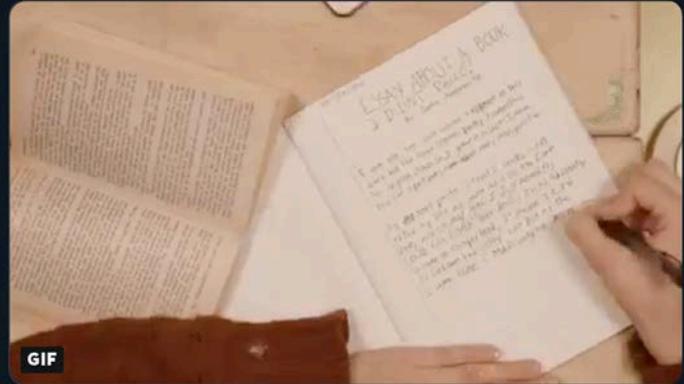
28 3 30



Speedy Paper @_speedypaper · Jan 4

Replying to @meggy97

Order any academic papers on speedypaper.com



28 3 30



Academic Researchers @AResearchers · Jan 4

Replying to @meggy97

follow back

28 3 30



COPYRIGHT INFRINGEMENT

- X Check the license model and give appropriate credit to authors for any external material.
- X Code
 - 54 LOC (out of 160 KLOC) of essential components may be sufficient to be considered a copyright violation.
 - American Court (Mertzel, 2008)
- X Papers, images, videos, icons, fonts, slide templates, ...



WHAT IS CITATION?

- X “Way you tell your readers that certain material in your work came from another source.” [plagiarism.org]
- X Why cite?
 - It’s the only way to use other people’s work without plagiarizing.
 - Helps readers understand and assess your work better.
 - Actually helps strengthen your work.
- X When to cite?
 - Whenever you use quotes, paraphrase, or use someone’s idea or work.



WHO SHOULD/SHOULD NOT BE AUTHORS ON YOUR PAPER?

ACM CRITERIA FOR AUTHORSHIP

- X Anyone listed as Author on an ACM manuscript submission must meet all the following criteria:
 - they have made substantial intellectual contributions to some components of the original work described in the manuscript
 - they have participated in drafting and/or revision of the manuscript
 - they are aware the manuscript has been submitted for publication
 - they agree to be held accountable for any issues relating to correctness or integrity of the work
- X Other contributors may be acknowledged at the end of the paper, before the bibliography.

EXAMPLE ACKNOWLEDGMENTS

ACKNOWLEDGMENTS

The authors thank Philip Guo, Max Goldman, Lauren Margulieux, Mark Guzdial, and the UID Group at MIT CSAIL for suggestions and feedback. This work was supported in part by the MIT EECS SuperUROP Program and Cisco Systems, Inc., by Quanta Computer through the Qmulus project, by NSF award SOCS-1111124, and by a Sloan Research Fellowship. Juho Kim is partly supported by the Samsung Fellowship.



THOUGHTS ON AUTHORSHIP

- X Have an explicit discussion, as early as possible.
- X When in doubt, talk to your advisor.
- X Just being your advisor doesn't automatically make them an author of your paper.
- X When it comes to borderline cases, I personally tend to err on the inclusive side.

THE AUTHOR LIST: GIVING CREDIT WHERE CREDIT IS DUE

The first author
Senior grad student on the project. Made the figures.

The third author
First year student who actually did the experiments, performed the analysis and wrote the whole paper. Thinks being third author is "fair".

The second-to-last author
Ambitious assistant professor or post-doc who instigated the paper.

Michaels, C., Lee, E. F., Sap, P. S., Nichols, S. T., Oliveira, L., Smith, B. S.

The second author
Grad student in the lab that has nothing to do with this project, but was included because he/she hung around the group meetings (usually for the food).

The middle authors
Author names nobody really reads. Reserved for undergrads and technical staff.

The last author
The head honcho. Hasn't even read the paper but, hey, he/she got the funding, and their famous name will get the paper accepted.

WWW.PHDCOMICS.COM



THE AUTHOR LIST

- X Depends on the discipline. Know the convention.
- X In CS,
 - Theoretical CS (math)
 - Alphabetical, equal contributions from all authors
 - Most other areas
 - The order is significant: order of decreasing contribution
 - First author (often a student) thought of as someone who did bulk of the work, thus gets bulk of the credit.
 - Faculty/senior typically listed last
- X It's often vague that reference letters need to explicitly explain the level of contribution.



THOUGHTS ON CREDIT & COLLABORATION

- X Don't ever hide other's contributions (e.g., getting a bit of [coding, study design, stats, writing, ...] help from others). Give appropriate credit.
- X Having more people involved does not necessarily undermine your own contribution.
- X Don't sacrifice the vision due to your limited skills and knowledge; find collaborators with whom you can achieve the vision together.
- X Open peer discussion & collaboration is one of the most effective ways to improve your idea.



WHAT DOES “ORIGINAL” RESEARCH MEAN?

- **Be original:** The content of submissions to NAACL-HLT 2019 (the ideas, the findings, the results and the words) should be original; that is, should not have been published (or be accepted for publication) in another refereed, archival form (such as a book, a journal, or a conference proceedings). Authors are referred to the ACL author guidelines for additional information on what constitutes existing publication.

Authors may present preliminary versions of their work in other venues that are not refereed and/or not archival (e.g. course reports, theses, non-archival workshops, or on preprint servers such as arXiv.org). Authors should list all such previous presentations in the submission form. This will help the area chairs if questions of originality arise.



WHAT ABOUT ARXIV?

1. Can I submit a paper to IJCAI if it already on arXiv?

Yes. We require that any submission to IJCAI must not be under review at any other *refereed* publication venue including archival conferences and journals. Papers put on arXiv do not violate this rule as long as they are not refereed.

2. A preliminary version of my paper has already been presented at a workshop or a conference without formal proceedings. It is ok to submit it to IJCAI?

Yes. This applies more generally to papers presented in any non-archival venue (that is, as long as there are no formally published proceedings with a publisher and an ISSN).

ETHICS IN DEALING WITH DATA

“I FOOLED MILLIONS INTO THINKING CHOCOLATE HELPS WEIGHT LOSS. HERE'S HOW.”

- X Journalistic experiment for how (intentionally) terrible & meaningless science can fool millions
- X Actually recruited people and ran a study
- X Findings: 10% faster weight loss with chocolate! (w/ p-hacking)
- X Submitted to 20 journals, acceptance within 24 hours
 - “outstanding manuscript,” and that for just 600 Euros it “could be accepted directly in our premier journal.”

Chocolate with high Cocoa content as a weight-loss accelerator

ORIGINAL

Johannes Bohannon¹,
Diana Koch¹,
Peter Homm¹,
Alexander Driehaus¹

¹ Institute of Diet and Health, Poststr. 37,
55126 Mainz, GERMANY

Abstract

Background: Although the focus of scientific studies on the beneficial properties of chocolate with a high cocoa content has increased in recent years, studies determining its importance for weight regulation, in particular within the context of a controlled dietary measure, have rarely been conducted.

Contact information:

✉ johannes@instituteofdiet.com.

THEN THE MEDIA RESPONDED

Pass the Easter Egg! New study reveals that eating chocolate doesn't affect your Body Mass Index ... and can even help you LOSE weight!

- New research from Roy Morgan reveals there's no proof that chocolate consumption affects BMI
- Currently two thirds of Australians eat chocolate at least once a month
- A study from German researchers has also found there's a connection between cocoa diets and increased weight loss
- Chocolate also found to benefit brain, heart and stress levels

By SAM BAILEY FOR DAILY MAIL AUSTRALIA



03.04 02:35 MIGnews.com

Шоколад - лучшая диета

WHAT MADE IT WORK?

- X P-hacking, poor science
- X Pay-to-publish journals
- X Poor peer review system
- X Lazy journalism
- X Our craving for catchy new research findings on diet

P-HACKING

- X Fiddle with study design & data until you get $p < 0.05$
 - Repeating the experiment, adding/removing/editing data
 - 18 different measurements from 15 people
- X Other dark patterns
 - Cherry-picking
 - Data manipulation
 - Other “honest” mistakes
- X “Why Most Published Research Findings Are False”, Ioannidis, 2005, PLOS Medicine.

HOW SCIENTISTS FOOL THEMSELVES — AND HOW THEY CAN STOP

Humans are remarkably good at self-deception. But growing concern about reproducibility is driving many researchers to seek ways to fight their own worst instincts.

COGNITIVE FALLACIES IN RESEARCH



HYPOTHESIS MYOPIA

Collecting evidence to support a hypothesis, not looking for evidence against it, and ignoring other explanations.



TEXAS SHARPSHOOTER

Seizing on random patterns in the data and mistaking them for interesting findings.



ASYMMETRIC ATTENTION

Rigorously checking unexpected results, but giving expected ones a free pass.



JUST-SO STORYTELLING

Finding stories after the fact to rationalize whatever the results turn out to be.

DEBIASING TECHNIQUES



DEVIL'S ADVOCACY

Explicitly consider alternative hypotheses — then test them out head-to-head.



PRE-COMMITMENT

Publicly declare a data collection and analysis plan before starting the study.



TEAM OF RIVALS

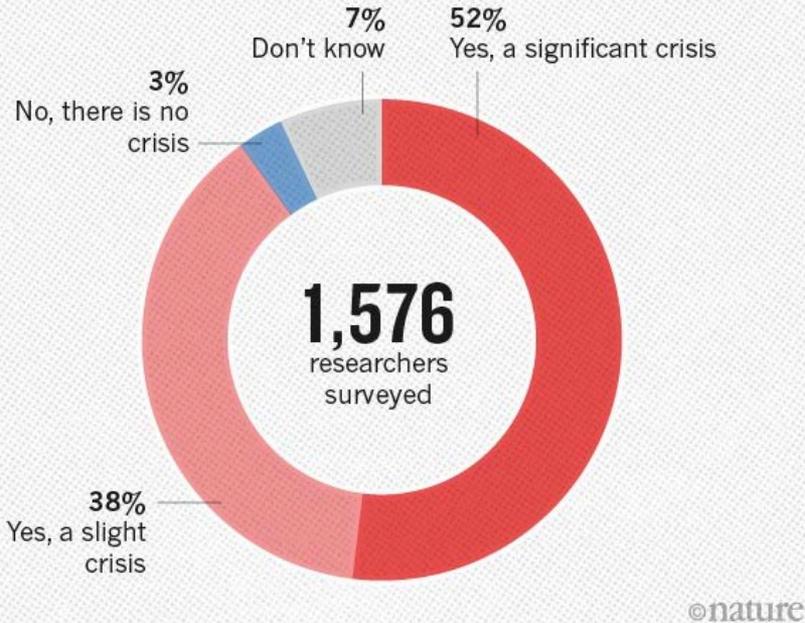
Invite your academic adversaries to collaborate with you on a study.



BLIND DATA ANALYSIS

Analyse data that look real but are not exactly what you collected — and then lift the blind.

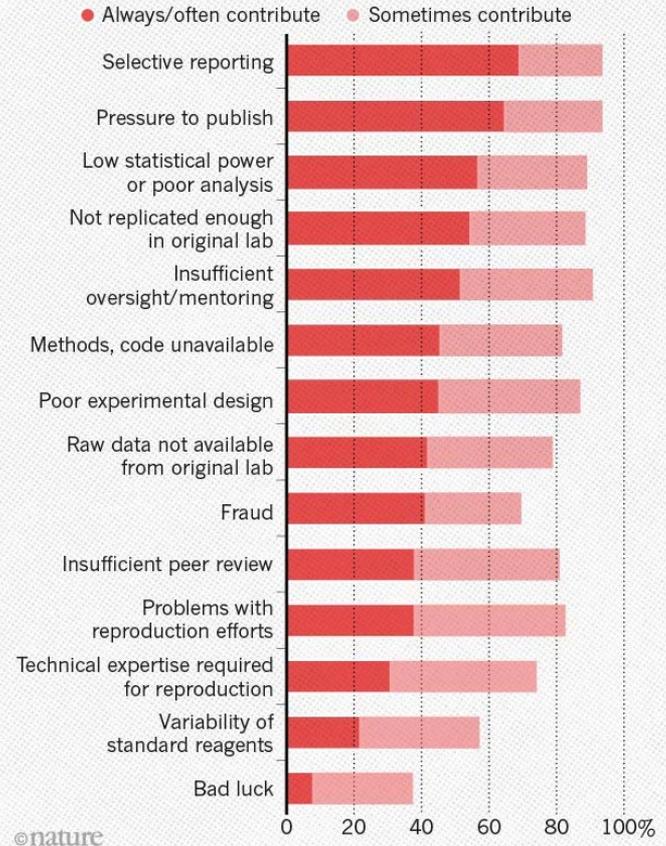
IS THERE A REPRODUCIBILITY CRISIS?



“More than 70% of researchers have tried and failed to reproduce another scientist's experiments, and more than half have failed to reproduce their own experiments.”

WHAT FACTORS CONTRIBUTE TO IRREPRODUCIBLE RESEARCH?

Many top-rated factors relate to intense competition and time pressure.



ETHICS IN RESEARCH WITH HUMAN SUBJECTS

Mostly drawn from
<https://www.kixlab.org/courses/cs374-spring-2018/classes/11-User-Testing/>



TUSKEGEE SYPHILIS EXPERIMENT

- X U.S. government studied the effects of untreated syphilis in African-American men in the rural South, under the guise of free health care
- X Not informed they had syphilis
- X Not treated even as proven, effective treatments like penicillin became available.
- X 6-month study → 40 years (1932–1972)



National Archives Atlanta, GA (U.S. government)



BELMONT REPORT:

ETHICAL GUIDELINES FOR HUMAN SUBJECT STUDIES

- X Respect for persons
 - voluntary participation & informed consent
 - protection of vulnerable populations (children, prisoners, people with disabilities, esp. cognitive)
- X Beneficence
 - do no harm
 - risks vs. benefits: risks to subjects should be commensurate with benefits of the work to the subjects or society
- X Justice
 - fair selection of subjects



INSTITUTIONAL REVIEW BOARD (IRB)

- X Research with people is subject to scrutiny
 - Most research institutions have an IRB that approves research-related user tests
 - KAIST has its own IRB. Review meetings held ~6 times a year.
- X IRB oversight is confined to research
 - “Research” is work leading to generalizable knowledge
 - “Practice” (clinical medicine, product development, class projects) does not require IRB approval
 - But all work with human beings should follow the IRB ethical guidelines, even if it doesn’t need IRB paperwork

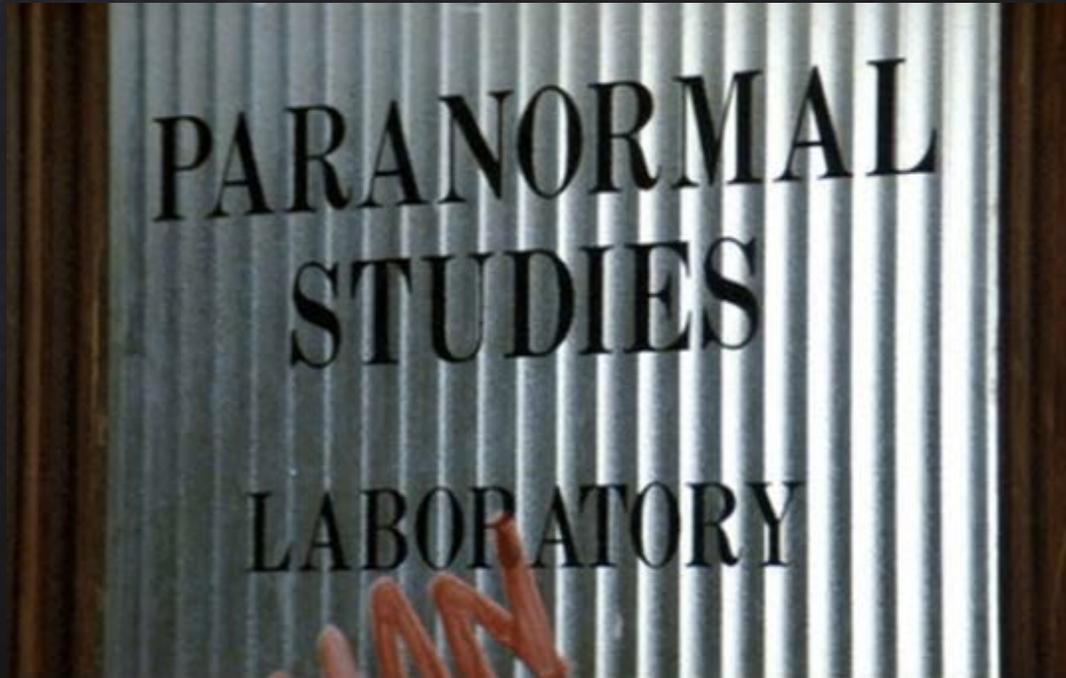


IRB APPROVAL

- X Human subjects training for all researchers
- X Main report
 - Objective
 - Descriptions of the system being tested
 - Task environment & material
 - Participants (minor, disabilities)
 - Methodology (deception study)
 - Tasks (cognitive, physical, emotional overhead)
 - Test measures (personal info)
- X Seems tedious but helps debug your study



ANY ETHICAL CONCERNS IN THIS STUDY?





TREAT THE USER WITH RESPECT

- X Time
 - Don't waste it
- X Comfort
 - Make the user comfortable
- X Informed consent
 - Inform the user as fully as possible
- X Privacy
 - Preserve the user's privacy
- X Control
 - The user can stop at any time

ETHICS IN COMPUTING RESEARCH

FACEBOOK EMOTION STUDY

“We show, via a massive (N = 689,003) experiment on Facebook, that emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness.”

Kramer et al., PNAS vol. 111 no. 24

- X A week in 2012
- X Modified News Feed to show more/less emotional posts
- X Seeing more negative posts → more negative posts
- X Seeing more positive posts → more positive posts
- X Seeing less emotional posts → less emotional posts

CAMBRIDGE ANALYTICA

X As many as 87M users' info was obtained, mined, sold, and used for the Trump campaign.



ETHICAL ISSUES IN LARGE-SCALE FIELD EXPERIMENTS

- X Informed consent? “Buried” in Terms of Service?
- X What does it mean to do large-scale user testing “in the wild”?
- X What about other A/B tests?
 - Active learning vs Lecture
 - Predicting illness based on search query

RESEARCH TOPICS WITH DIRECT ETHICAL IMPLICATIONS

- X algorithmic bias
- X censorship
- X privacy / security
- X anonymity / harassment
- X misinformation / fake news
- X surveillance
- X terms and conditions
- X information inequity / accessibility

“PROGRAMMER”

Original:



Balancing gender:



Balancing skin color:

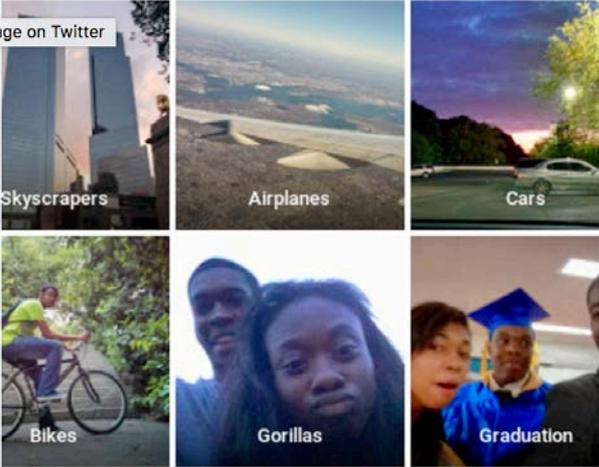


Balancing age:



IMPLICIT RACIAL BIAS

View image on Twitter



Skyscrapers Airplanes Cars

Bikes Gorillas Graduation

 **jackyalciné**
@jackyalcine [Follow](#)

Google Photos, y'all fucked up. My friend's not a gorilla.

10:22 AM - 29 Jun 2015

← ↻ 3,186 ❤️ 1,999

INADVERTENT ALGORITHMIC CRUELTY

X *“Yes, my year looked like that. True enough. My year looked like the now-absent face of my little girl. It was still unkind to remind me so forcefully,”*





MITIGATING THE NEGATIVE IMPACTS OF COMPUTING

- X ACM's Future of Computing Academy's proposal
- X *"Peer reviewers should require that papers and proposals rigorously consider all reasonable broader impacts, both positive and negative."*
- X Examples
 - Crowdsourcing: minimum wage
 - Blockchain: energy usage & corresponding carbon footprint
 - Lip reading technology: surveillance
 - Algorithmic decision making: fairness, accountability, transparency



NEURIPS 2020 CFP

4. Authors need to declare if a previous version of their submission was rejected at any peer-reviewed venue within the past 12 months, and, if so, summarize the changes to the current version. This information should be entered into CMT during the submission process.
5. In order to provide a balanced perspective, authors are required to include a statement of the potential broader impact of their work, including its ethical aspects and future societal consequences. Authors should take care to discuss both positive and negative outcomes.
6. Authors are required to provide an explicit disclosure of funding (financial activities supporting the submitted work) and competing interests (related financial activities outside the submitted work) that could result in conflicts of interest. This section should be added to the camera-ready version of accepted papers. More information can be found [here](#).
7. We strongly encourage (but do not require) accompanying code and data to be submitted with accepted papers that contribute and present experiments with a new algorithm or new dataset. Moreover, we encourage authors to upload their code as part of their supplementary material at submission time in order to help reviewers assess the quality of the work. Check [the policy, guidelines and templates](#) for code submission for further details.



TAKE-HOME MESSAGES

- X Ethical issues apply to all phases of research.
 - Know what they are, learn to properly address them.

- X Ethical implications exist in all areas of computing research.
 - Know what they are, learn to properly address them.

ETHICAL GUIDELINES

- X ACM Code of Ethics (<https://www.acm.org/code-of-ethics>)
- X IEEE Code of Ethics
(<https://www.ieee.org/about/corporate/governance/p7-8.html>)
- X APA's Ethical Principles of Psychologists and Code of Conduct
(<https://www.apa.org/ethics/code/index.aspx>)
- X National Society of Professional Engineers Code of Ethics
(<https://www.nspe.org/resources/ethics/code-ethics>)