

Tips for Scientific Writing

Prof. Gabriela Alves Werb, Ph.D.

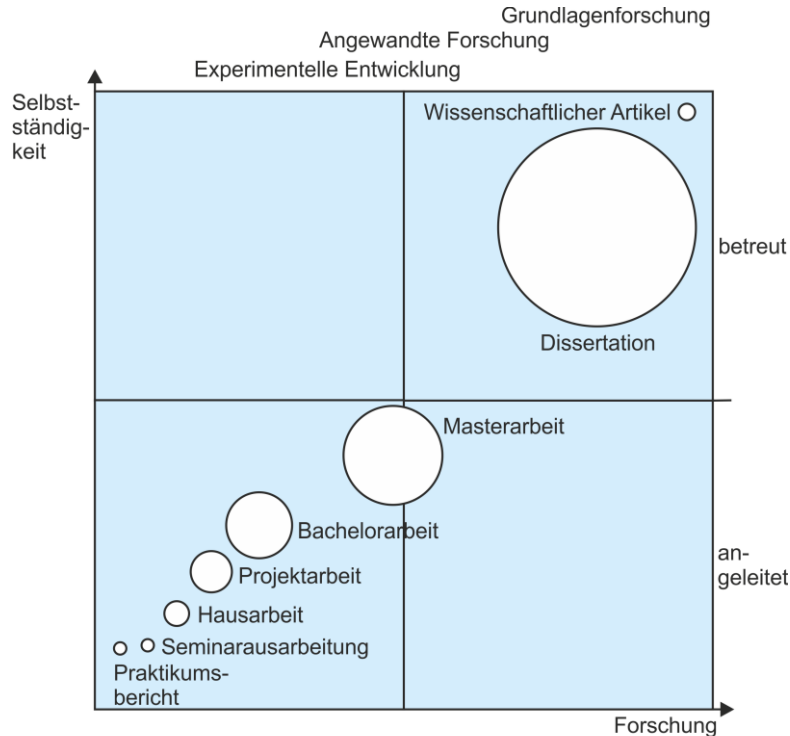
gabriela.alveswerb@fb2.fra-uas.de

Last Update: April 2023

Agenda

- 1 Research Plan**
- 2 Writing Tips
- 3 Literature Search
- 4 Outlining the Methods / Data
- 5 Abstract

Different Types of Scientific Work

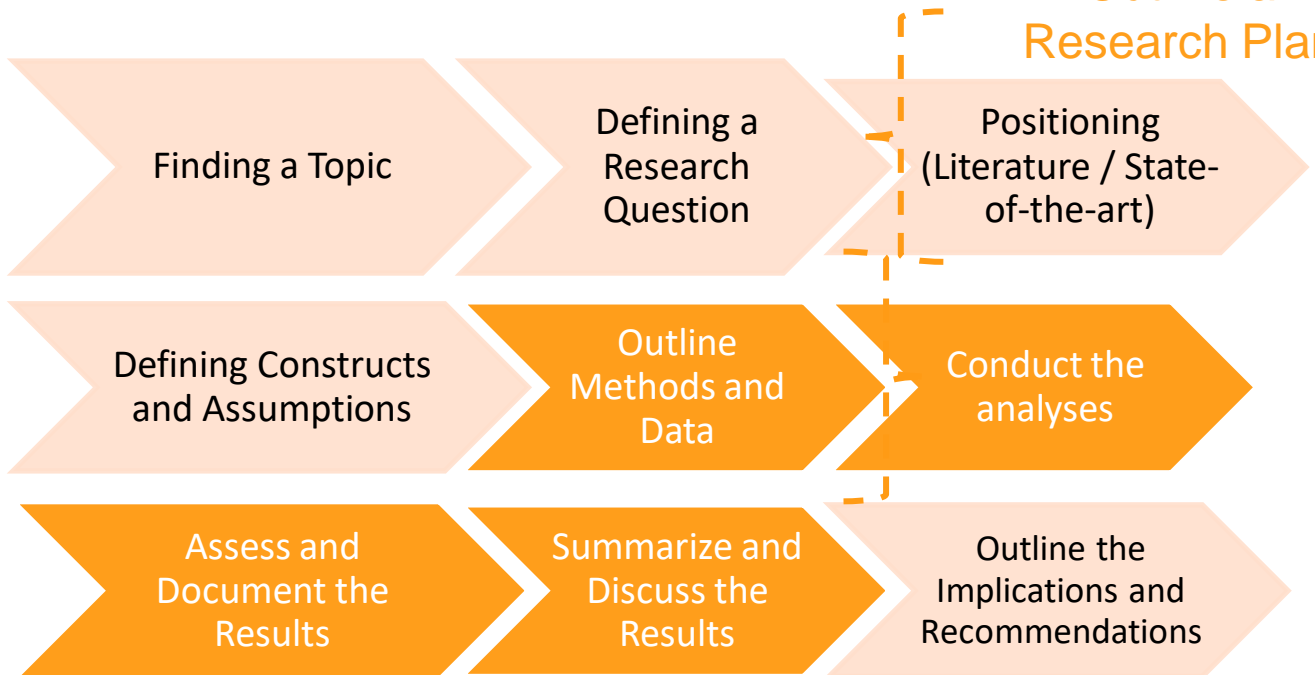


Balzert et. al (2011)

Legende: Die Kreisgröße symbolisiert den maximalen Textumfang der wissenschaftlichen Arbeit

The Steps of a Research Project / Thesis

Outline a Research Plan



Agenda

- 1 Research Plan
- 2 Writing Tips**
- 3 Literature Search
- 4 Outlining the Methods / Data
- 5 Abstract

The Writing Process

- Objective writing
- Avoid absolute statements, especially those without any corroboration (always, never, exclusively...)
- Define all acronyms and crucial concepts and terms
- If your problem is too complex or difficult to explain, use an example or visualization (fictitious company, small numerical example)

Telling a Story

- Think of ways to make your story appealing and interesting
- What are the 3 components of a story?
 - Setting and Characters (once upon a time...)
 - Conflict (what is the problem that needs to be solved?)
 - Resolution (the problem is fixed!)

Engage Your Reader

- Try to **engage the reader** until the very end
- Don't go all the way back to "Adam and Eve"
- Try to **avoid unnecessary details**
- Use **examples** if the concepts are too complicated or abstract

Title Choice

- The title should grasp others' **interest**
- It can (but must not be) funny or intriguing
- Avoid very long titles

- Pursuing the early voter: Does the early bird get the worm?
- On Storks and Babies: Correlation, Causality and Field Experiments
- What Makes Online Content Viral?

vs

- Mapping Time: How the Spatial Representation of Time Influences Intertemporal Choices
- An Experimental Investigation of the Effects of Retargeted Advertising: The Role of Frequency and Timing

Good Practices

- **Avoid** using **passive voice** whenever possible
- **Avoid long nested** sentences (... , but ... , which ...)
- Try to **describe precisely** what you want to do
- Use **transitions** to connect your sentences and paragraphs
- **Proofread, proofread, proofread** (Word spellchecker, Grammarly, Duden online)

Does Your Story Really Fit Together?

Read everything backwards (from Conclusion to Introduction)

- See **whether quality “declines”** throughout the article

What did you state in the beginning as **goals**?

- Do you really achieve them?
- Do they fit with the methods you use and results you obtain?

Are your **labels cohesive**?

- E.g., firms, companies, corporations – choose **only one** and use it everywhere

Agenda

- 1 Research Plan
- 2 Writing Tips
- 3 Literature Search**
- 4 Outlining the Methods / Data
- 5 Abstract

Getting Started – Literature Research

3 W's and 2 H's

- **What** am I searching for?
- **Where** do I search for it?
- **How** do I search for it?
- **What** do I use?
- **How** do I obtain it?

Different Types of Scientific Sources

Scientific Journals

Conference
Proceedings

Monographs

Internet Sources
(use with care!)

Working Papers /
Work on Pre-Print
Servers
(use with care...)

Law / Regulation
Texts

Norms and
Standards (e.g.,
ISO, COBIT, ITIL, ...)

Court Rulings
(perhaps not so
relevant in our
field)

Requirements

Primary and other sources

Differentiate between primary (first-hand evidence) and secondary sources



The source is...

understandable
published
identifiable
verifiable



Is it worthy of citing?

Does it adhere to scientific standards?

Does it involve a serious, academically recognized institution?



Balzert et. al (2011)

Systematic Overview

- Previously define a set of inclusion **criteria**
 - Which research **streams / theories** are **relevant**?
 - Do I need to rely on previously established **constructs**?
 - Tip: Typically, there are **many papers** talking about a given topic – focus on the seminal articles / **most influential ones**
- Critically assess the source's **quality**
 - Does it “make the cut” of the quality criteria?



Systematic Overview

- Systematically extract the **relevant information**
 - Which key information do I need?
 - It is a concept?
 - It is a set of results – e.g., this and this research shows that X affects Y?
- **Summarize and present** the information
 - E.g., different paragraphs for each set of results, when there is controversy
 - Table comparing a set of studies similar to yours



Search Strategy

- Specify a set of relevant keywords
- Identify similar keywords / synonyms
- Sometimes: constrain the search to specific journals
 - E.g., “keyword 1 keyword 2 source: XXX”



Search Databases

- **Google Scholar**
- **Semantic Scholar**
- IEEE Xplore
- ACM Digital Library
- SpringerLink
- Science Direct (Elsevier)
- Wiley Online Library
- JSTOR
- ...



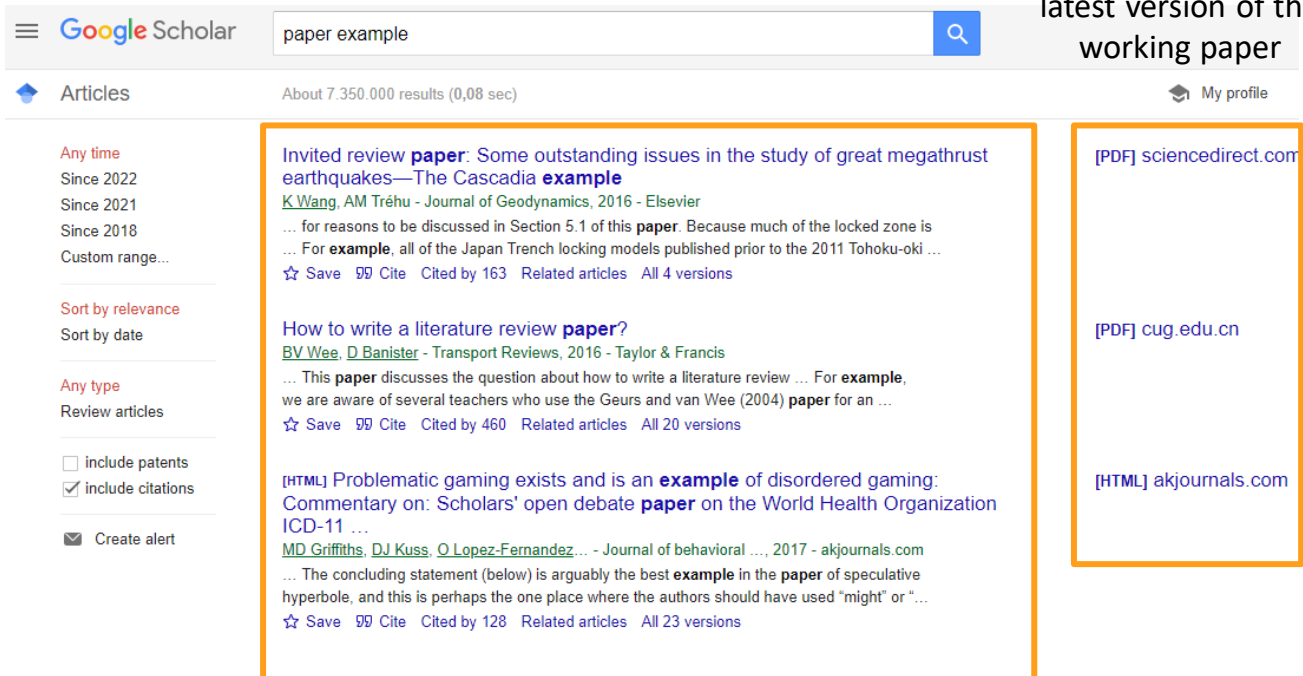
These compile more
sources



Getting Access (I)

These links are typically behind paywalls (unless open access)

These links typically offer direct access, might be the author's copy or the latest version of the working paper



The screenshot shows a Google Scholar search for 'paper example' with approximately 7,350,000 results. The left sidebar contains filters for 'Any time' (Since 2022, 2021, 2018, Custom range...), 'Sort by relevance' (Sort by date), and 'Any type' (Review articles, include patents, include citations, Create alert). The main results are:

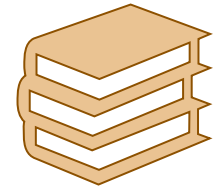
- Invited review paper:** Some outstanding issues in the study of great megathrust earthquakes—The Cascadia **example**
[K Wang, AM Tréhu](#) - Journal of Geodynamics, 2016 - Elsevier
 ... for reasons to be discussed in Section 5.1 of this **paper**. Because much of the locked zone is ... For **example**, all of the Japan Trench locking models published prior to the 2011 Tohoku-oki ...
 ☆ Save 📄 Cite Cited by 163 Related articles All 4 versions
- How to write a literature review paper?**
[BV Wee, D Banister](#) - Transport Reviews, 2016 - Taylor & Francis
 ... This **paper** discusses the question about how to write a literature review ... For **example**, we are aware of several teachers who use the Geurs and van Wee (2004) **paper** for an ...
 ☆ Save 📄 Cite Cited by 460 Related articles All 20 versions
- [HTML] Problematic gaming exists and is an example of disordered gaming: Commentary on: Scholars' open debate paper on the World Health Organization ICD-11 ...**
[MD Griffiths, DJ Kuss, O Lopez-Fernandez](#) ... - Journal of behavioral ..., 2017 - akjournals.com
 ... The concluding statement (below) is arguably the best **example** in the **paper** of speculative hyperbole, and this is perhaps the one place where the authors should have used "might" or "...
 ☆ Save 📄 Cite Cited by 128 Related articles All 23 versions

On the right side, three links are highlighted in orange boxes:

- [PDF] sciencedirect.com
- [PDF] cug.edu.cn
- [HTML] akjournals.com

Getting Access (II)

- FRA-UAS's Library – “Katalog Plus”
 - <https://hds.hebis.de/fuas/index.php>
- Remote access (e.g., from home) using VPN
 - <https://www.frankfurt-university.de/de/hochschule/bibliothek/fernzugriff/>
- Search for the specific title/authors in working paper or pre-print repositories
 - ResearchGate
 - SSRN
 - arXiv
 - ...



Check **Every Reference** in the Reference List

Typical problems:

- References without publication **year**
- **Online** references without **link** or **access date**
- A few published articles with **DOI**, others without
- **Published** articles referencing the **working paper repository** (e.g., research gate, SSRN) instead of the journal
- A few article **names capitalized** or in **caps lock**, others not

Agenda

- 1 Research Plan
- 2 Writing Tips
- 3 Literature Search
- 4 Outlining the Methods / Data**
- 5 Abstract

Inductive vs. Deductive Research

Inductive (bottom-up)

Tentative
Hypothesis

Pattern

Observations

Theory

Deductive (top-down)

Hypothesis

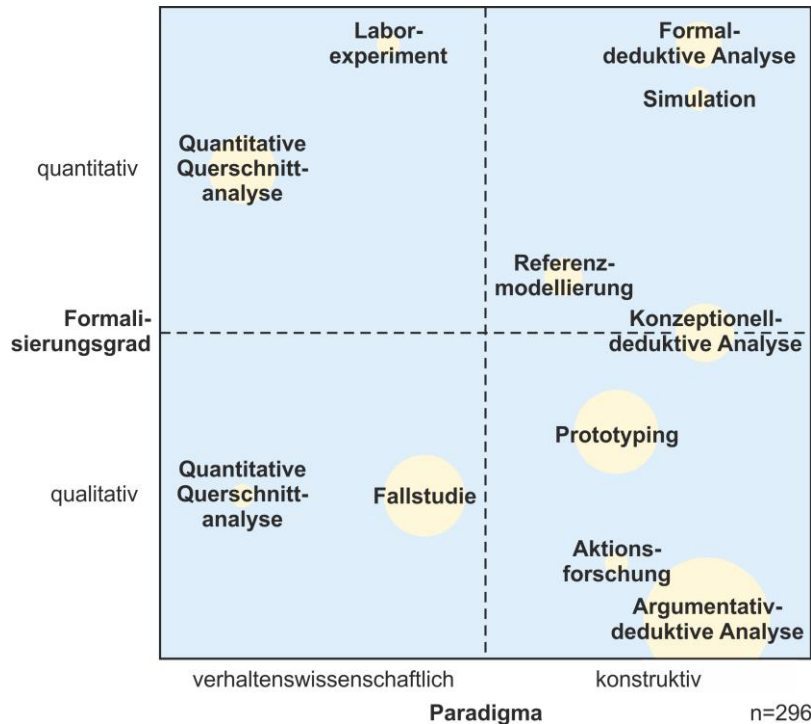
Prediction

Observations

Confirmation
/ Rejection

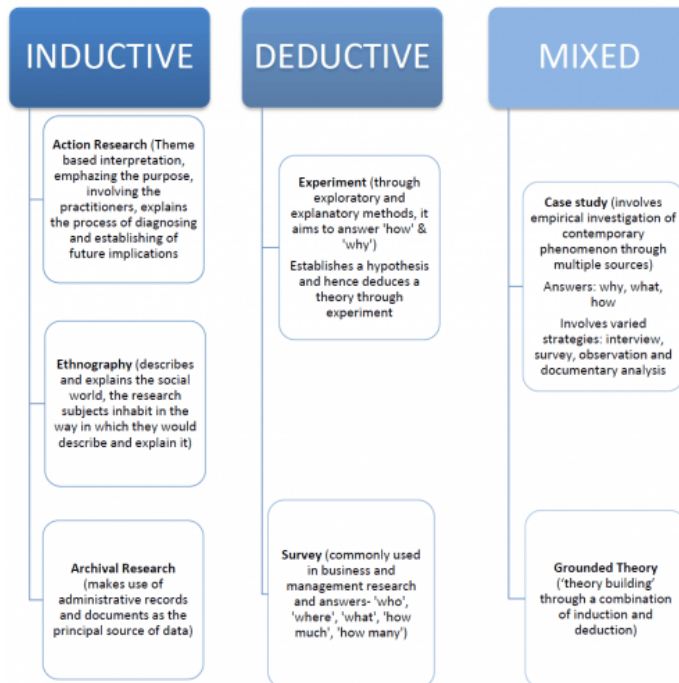
Balzert et. al (2011)

Research Methods in Business Informatics



Balzert et. al (2011)

Mixed Methods



Datt, S. and Chetty, P. (2016)

Popular Sources of Data

- Experiments
- Observational Data
 - Policy changes
 - Natural experiments
- Surveys
- Structured Interviews
-

Agenda

- 1 Research Plan
- 2 Writing Tips
- 3 Literature Search
- 4 Outlining the Methods / Data
- 5 Abstract**

Abstract

Self-Management in heterogeneous Networks using a Service-Oriented Architecture

<p>Claudia Baltes Regionales Hochschullehrerzentrum Kaiserslautern University of Kaiserslautern Paul-Ehrlich-Strasse, Gebäude 34 67663 Kaiserslautern, Germany Email: baltes@rhzk.uni-kl.de</p>	<p>Patrick Koppen Fundamental Generic Networking GmbH Research and Development Gottlieb-Daenke-Strasse, Gebäude 38 67663 Kaiserslautern, Germany Email: koppen@g-networking.de</p>	<p>Paul Müller AG IC3Y University of Kaiserslautern Fachbereich Informatik Paul-Ehrlich-Strasse, Gebäude 34 67663 Kaiserslautern, Germany Email: p.mueller@informatik.uni-kl.de</p>
---	--	---

Abstract.—Self-Managing Networks promise to help network administrators in their daily work and are less error-prone. This paper introduces an architecture for monitoring a local area network and protecting it against misuse. A regulatory system is automatically generated or restricted a malicious host from the network is introduced. Whereby the focus is to support even heterogeneous networks. Various sensors like an IDS or accounting system may make the decision that a host should be detected or rate-limited. A request is sent to the regulatory system and the problem is an automatically determines the best way to enforce this regulation as there are various possibilities e.g. disabling a port or applying packet filters. Models for locating the host and configuring the network devices complete the presented solution.

A. Related Work

Formerly hosts within a LAN were usually considered to be untrusted. This firewall at the border of administrative networks domains were used to protect local hosts against threats from the outside. But it turned out that there must also be protection against malicious local hosts. The first step was to develop mechanisms to detect such malicious behavior, which are commonly distinguished by the method of detection (signature or anomaly based) and by the location of the sensors (network or host). Intrusion Detection Systems (IDS) usually combine several detector mechanisms [2], [14], [15]. Many commercial and free IDS exist and are widely used today, e.g. [6]. Using an IDS still requires an administrator to react on detected potential threats. So called Intrusion Prevention Systems (IPS) or Intrusion Response Systems (IRS) [7] promise to exceed IDS by actively taking countermeasures, for example rate-limitation [8] or detaching a host from the network. Commercial IRS and IPS are available but these are designed for homogeneous networks and often rely on vendor specific functionality.

In practice there are often several very different incidents that require different kinds of reaction by an administrator. For example a host utilizing too much bandwidth might trigger different reactions than a host which attacks other systems. Further in heterogeneous networks it is often necessary to use different countermeasures due to different capabilities of intermediate network devices. Thus our goal is to develop an open architecture which is adaptable to different network environments. Whereby a focus is on the determination of applicable countermeasures.

B. Challenge

Each network has its specific requirements according to what should be monitored and where to place the sensors. Thus the first challenge is handling several independent sensors of different types. All sensors may request to regulate a host or to revoke the regulation. So the regulatory system must deal with overlapping and even contradictory requests from the sensors.

The second challenge is for the regulatory system to determine if and how this request should be enforced. Depending

I. INTRODUCTION

The importance of networks is still growing, today almost all workplaces are equipped with computers using the network to access data and services. Enhancing these networks by integrating VoIP or the increasing number of applications that are based on services residing somewhere within the network makes the continuous availability of the network even more important. Proportional to the number of devices and requirements for features like high availability and QoS grows the complexity and administrative burden of managing the network. In the near future our networks will not be manageable without supportive tools.

One problem in today's networks is the constant exposure to security threats [1]. Viruses, worms, hackers, hacker attacks or misconfigured hosts must be detected and eliminated. There may be other malpractices e.g. exceeding a given quota. Any disturbing behavior must be detected and countermeasures must be taken. Typically a misbehaving host will be detached from the network or rate-limited. This paper introduces an approach for self-managing heterogeneous networks which will support a network administrator in dealing with malfunctions or misbehaving hosts. The basic approach is to build a regulatory system which is able to use several different sensors and actuators. Extensibility and adaptability – which is crucial for heterogeneous environments – is achieved using a service-oriented architecture.

What is an Abstract?

- Short summary of your paper
- Approx. 150 to 250 words
- Start with a longer text and refine / shorten it

What Makes a Good Abstract?

- Short
- Mostly in present tense
- Precise and objective language (no abstract statements or very long sentences)
- Clear connections between the sentences (no sequence of loosely related
- Goal: spike the interest of readers (is this paper relevant for me?)
- No citations or references to tables, figures or other elements of the main document

Construction



Context / Background



Goal and Research Question(s)



Methods / Data



Results and Implications

Example I (234 Words)

Context

Firms can now offer personalized recommendations to consumers who return to their website, using consumers' previous browsing history on that website. In addition, online advertising has greatly improved in its use of external browsing data to target Internet ads. Dynamic retargeting integrates these two advances by using information from the browsing history on the firm's website to improve advertising content on external websites. When surfing the Internet, consumers who previously viewed products on the firm's website are shown ads with images of those same products. To examine whether this is more effective than simply showing generic brand ads, the authors use data from a field experiment conducted by an online travel firm. Surprisingly, the data suggest that dynamic retargeted ads are, on average, less effective than their generic equivalents. However, when consumers exhibit browsing behavior that suggests their product preferences have evolved (e.g., visiting review websites), dynamic retargeted ads no longer underperform. One explanation for this finding is that when consumers begin a product search, their preferences are initially construed at a high level. As a result, they respond best to higher-level product information. Only when they have narrowly construed preferences do they respond positively to ads that display detailed product information. This finding suggests that in evaluating how best to reach consumers through ads, managers should be aware of the multistage nature of consumers' decision processes and vary advertising content along these stages.

<https://doi.org/10.1509/jmr.11.0503>

Example I (234 Words)

Research Question

Firms can now offer personalized recommendations to consumers who return to their website, using consumers' previous browsing history on that website. In addition, online advertising has greatly improved in its use of external browsing data to target Internet ads. Dynamic retargeting integrates these two advances by using information from the browsing history on the firm's website to improve advertising content on external websites. **When surfing the Internet, consumers who previously viewed products on the firm's website are shown ads with images of those same products. To examine whether this is more effective than simply showing generic brand ads,** the authors use data from a field experiment conducted by an online travel firm. Surprisingly, the data suggest that dynamic retargeted ads are, on average, less effective than their generic equivalents. However, when consumers exhibit browsing behavior that suggests their product preferences have evolved (e.g., visiting review websites), dynamic retargeted ads no longer underperform. One explanation for this finding is that when consumers begin a product search, their preferences are initially construed at a high level. As a result, they respond best to higher-level product information. Only when they have narrowly construed preferences do they respond positively to ads that display detailed product information. This finding suggests that in evaluating how best to reach consumers through ads, managers should be aware of the multistage nature of consumers' decision processes and vary advertising content along these stages.

<https://doi.org/10.1509/jmr.11.0503>

Example I (234 Words)

Methods

Firms can now offer personalized recommendations to consumers who return to their website, using consumers' previous browsing history on that website. In addition, online advertising has greatly improved in its use of external browsing data to target Internet ads. Dynamic retargeting integrates these two advances by using information from the browsing history on the firm's website to improve advertising content on external websites. When surfing the Internet, consumers who previously viewed products on the firm's website are shown ads with images of those same products. To examine whether this is more effective than simply showing generic brand ads, **the authors use data from a field experiment conducted by an online travel firm.** Surprisingly, the data suggest that dynamic retargeted ads are, on average, less effective than their generic equivalents. However, when consumers exhibit browsing behavior that suggests their product preferences have evolved (e.g., visiting review websites), dynamic retargeted ads no longer underperform. One explanation for this finding is that when consumers begin a product search, their preferences are initially construed at a high level. As a result, they respond best to higher-level product information. Only when they have narrowly construed preferences do they respond positively to ads that display detailed product information. This finding suggests that in evaluating how best to reach consumers through ads, managers should be aware of the multistage nature of consumers' decision processes and vary advertising content along these stages.

<https://doi.org/10.1509/jmr.11.0503>

Example I (234 Words)

Results

Firms can now offer personalized recommendations to consumers who return to their website, using consumers' previous browsing history on that website. In addition, online advertising has greatly improved in its use of external browsing data to target Internet ads. Dynamic retargeting integrates these two advances by using information from the browsing history on the firm's website to improve advertising content on external websites. When surfing the Internet, consumers who previously viewed products on the firm's website are shown ads with images of those same products. To examine whether this is more effective than simply showing generic brand ads, the authors use data from a field experiment conducted by an online travel firm. Surprisingly, the data suggest that dynamic retargeted ads are, on average, less effective than their generic equivalents. However, when consumers exhibit browsing behavior that suggests their product preferences have evolved (e.g., visiting review websites), dynamic retargeted ads no longer underperform. One explanation for this finding is that when consumers begin a product search, their preferences are initially construed at a high level. As a result, they respond best to higher-level product information. Only when they have narrowly construed preferences do they respond positively to ads that display detailed product information. This finding suggests that in evaluating how best to reach consumers through ads, managers should be aware of the multistage nature of consumers' decision processes and vary advertising content along these stages.

<https://doi.org/10.1509/jmr.11.0503>

Example I (234 Words)

Firms can now offer personalized recommendations to consumers who return to their website, using consumers' previous browsing history on that website. In addition, online advertising has greatly improved in its use of external browsing data to target Internet ads. Dynamic retargeting integrates these two advances by using information from the browsing history on the firm's website to improve advertising content on external websites. When surfing the Internet, consumers who previously viewed products on the firm's website are shown ads with images of those same products. To examine whether this is more effective than simply showing generic brand ads, the authors use data from a field experiment conducted by an online travel firm. Surprisingly, the data suggest that dynamic retargeted ads are, on average, less effective than their generic equivalents. However, when consumers exhibit browsing behavior that suggests their product preferences have evolved (e.g., visiting review websites), dynamic retargeted ads no longer underperform. One explanation for this finding is that when consumers begin a product search, their preferences are initially construed at a high level. As a result, they respond best to higher-level product information. Only when they have narrowly construed preferences do they respond positively to ads that display detailed product information. **This finding suggests that in evaluating how best to reach consumers through ads, managers should be aware of the multistage nature of consumers' decision processes and vary advertising content along these stages.**

Implications

<https://doi.org/10.1509/jmr.11.0503>

Example II (102 Words)

Context + Methods

We explore data from a field test of how an algorithm delivered ads promoting job opportunities in the science, technology, engineering and math fields. This ad was explicitly intended to be gender neutral in its delivery. Empirically, however, fewer women saw the ad than men. This happened because younger women are a prized demographic and are more expensive to show ads to. An algorithm that simply optimizes cost-effectiveness in ad delivery will deliver ads that were intended to be gender neutral in an apparently discriminatory way, because of crowding out. We show that this empirical regularity extends to other major digital platforms.

<https://doi.org/10.1287/mnsc.2018.3093>

Example II (102 Words)

Research Question

We explore data from a field test of how an algorithm delivered ads promoting job opportunities in the science, technology, engineering and math fields. **This ad was explicitly intended to be gender neutral in its delivery. Empirically, however, fewer women saw the ad than men.** This happened because younger women are a prized demographic and are more expensive to show ads to. An algorithm that simply optimizes cost-effectiveness in ad delivery will deliver ads that were intended to be gender neutral in an apparently discriminatory way, because of crowding out. We show that this empirical regularity extends to other major digital platforms.

<https://doi.org/10.1287/mnsc.2018.3093>

Example II (102 Words)

We explore data from a field test of how an algorithm delivered ads promoting job opportunities in the science, technology, engineering and math fields. This ad was explicitly intended to be gender neutral in its delivery. Empirically, however, fewer women saw the ad than men. **This happened because younger women are a prized demographic and are more expensive to show ads to. An algorithm that simply optimizes cost-effectiveness in ad delivery will deliver ads that were intended to be gender neutral in an apparently discriminatory way, because of crowding out. We show that this empirical regularity extends to other major digital platforms.**

Results + Implications

<https://doi.org/10.1287/mnsc.2018.3093>

References

- Balzert, H., Schröder, M., & Schäfer, C. (2011), *Wissenschaftliches Arbeiten-Ethik, Inhalt & Form wiss. Arbeiten, Handwerkszeug, Quellen, Projektmanagement*.
- Pollock, T. G., & Bono, J. E. (2013), *Being Scheherazade: The Importance of Storytelling in Academic Writing*.
- Tischler, M. E. (2005), *Scientific Writing Booklet, Department of Biochemistry & Molecular Biophysics University of Arizona*.
- Zinsser, W. (2006), *On Writing Well: The Classic Guide to Writing Nonfiction*. New York: HarperCollins.