

University of Groningen

## Computer programming skills: A cognitive perspective

Graafsma, Irene

DOI:  
[10.33612/diss.168003240](https://doi.org/10.33612/diss.168003240)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2021

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Graafsma, I. (2021). *Computer programming skills: A cognitive perspective*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.168003240>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# **Computer programming skills: A cognitive perspective**

Irene L. Graafsma



The research reported in this thesis has been carried out under the auspices of the Center for Language and Cognition Groningen (CLCG), the research school of Behavioral and Cognitive Neuroscience (BCN) of the University of Groningen, and the International Doctorate for Experimental Approaches to Language And Brain' (IDEALAB) of the Universities of Groningen (NL), Newcastle (UK), Potsdam (DE) and Macquarie University, Sydney (AU).

Publication of this thesis was financially supported by the Graduate School of Humanities (GSH) of the University of Groningen and by the research school of Behavioral and Cognitive Neuroscience (BCN) of the University of Groningen.



Groningen Dissertations in Linguistics 201

© 2021, Irene L. Graafsma

Cover: Ella Hope

Printed by Ipskamp Printing (NL) - [www.ipskampprinting.nl](http://www.ipskampprinting.nl)



# Computer programming skills: a cognitive perspective

## PhD thesis

to obtain the joint degree of PhD at the  
University of Groningen, University of Potsdam, Macquarie University and Newcastle  
University

on the authority of the  
Rector Magnificus of the University of Groningen Prof. C. Wijmenga,  
President of the University of Potsdam, Prof. O. Günther,  
Deputy Vice Chancellor of Macquarie University, Prof. S. Bruce Downton, and  
Vice Chancellor of Newcastle University, Prof. Ch. Day

and in accordance with  
the decision by the College of Deans.

This thesis will be defended in public on

Thursday 29 April 2021 at 14:30 hours

by

**Irene Lotte Graafsma**

born on 19 April 1993  
in Saint Martin D'Hères, France

**Supervisors**

Prof. Y.R.M. Bastiaanse

Prof. L. Nickels

**Co-supervisors**

Dr. E. Marinus

Dr. S. Robidoux

Dr. S. Popov

**Assessment Committee**

Prof. M.J. Guzdial

Prof. M. Nissim

Prof. G.J.M. van Noord

Prof. P.C.J. Segers





# Acknowledgements



This PhD has been a great academic and personal adventure. Being able to do research in a fascinating, relatively new field whilst living in two amazing countries has been a fantastic opportunity that would not have been possible without the extensive support of a large group of wonderful people.

First of all, I would like to thank my supervisors. I want to thank my promotor Roelien Bastiaanse for her help and support, especially with Chapter 5, and for her general help with the thesis writing and submission process. Her extensive experience has been very valuable in shaping this thesis. Special thanks go out to Eva Marinus, without whom the project would have never existed. She supervised me throughout the entire PhD, and has read my drafts many, many times. Although we were only in the same physical place briefly at the start of the PhD, Eva has stayed very involved and has inspired me with her passion for the subject. I also want to thank both Lyndsey Nickels and Serje Robidoux for their help and supervision during my time at Macquarie, and for their help with the analysis and writing once I was in Groningen. I also want to thank them for providing such a supportive and friendly atmosphere at Macquarie, where I very quickly felt at home. I also want to thank Serje for helping with the testing during the pilot study. I want to thank Nathan Caruana for supervising the writing of Chapter 4. Thanks to his thorough and always very timely feedback the writing process for this chapter was very smooth. I also learnt a lot, both about autism research, and about the writing and publishing process. Finally, I want to thank Srdjan Popov for his help and supervision of Chapter 5. I would never have been able to set up, run, and write up this experiment without his help.

I would also like to thank my assessment committee, Mark Guzdial, Malvina Nissim, Gertjan van Noord and Eliane Segers. Thank you for reading my thesis in such detail, and for your helpful comments.

I also want to thank Lesley McKnight and Katie Webb at Macquarie University, and Alice Pomstra, Marijke Wubbolts and Christina Englert at the University of Groningen for all their help with organisational matters.

I would also like to thank the IDEALAB program in general, for making this beautiful international PhD program possible. Specifically, I would like to thank the directors, David Howard, Barbara Hole, Lyndsey Nickels, Roelien Bastiaanse and Gabriele Miceli. Thank you

for the amazing program, and in particular for the summer and winter schools, where I learnt a lot from your questions and expertise. I would also like to thank the other IDEALAB PhD students for their excellent questions and feedback during the winter and summer schools and for making the moves across countries more fun.

Throughout the PhD I have had help from several other kind individuals. Special thanks go out to Matthew Roberts. Thanks to Matt we were able to test all students from the programming units in the Department of Computing at Macquarie. Matt even arranged for testing to be integrated in the course tutorials. The first three chapters of the thesis would not have been possible without his help and collaboration. Matt also contributed to the design of the studies and was our programming and computing expert in the Cognition of Coding research group.

I would also like to thank the other two members of the Cognition of Coding group, Vince Polito and Judy Zhu. They contributed many good ideas, as well as important help and feedback with test selection and design. They also contributed to the writing and publication process of Chapter 3.

I specifically want to thank those who helped with recruitment of participants for the EEG study: Jan van Beek, Rik Teerling and Bart Barnard. More generally, I also want to thank the department of Computational Linguistics at the University of Groningen, as well as the student interns in the Department of Computing at Macquarie University who helped design the EEG stimuli and helped with the pilots and test digitisation for Chapters 2, 3 and 4.

I want to thank Ella Hope for designing the beautiful cover of this book. She immediately understood what I was looking for, made an amazing design and helped me to get it printed correctly.

Also, I am very grateful to all the (former and present) members of the Neurolinguistics research group (not an exhaustive list): Adrià, Aida, Annie, Atilla, Ben, Dörte, Frank, Jakolien, Jidde, Juliana, Kaimook, Liset, Nathaniel, Nermina, Pauline, Roel, Roelant, Roelien, Sara, Seçkin, Serine, Srdan, Suzan, Svetlana, Teja, Vânia, and Yulia. Thank you for the interesting meetings, and for your help and feedback on my presentations. In particular I want to thank Seçkin for teaching me to do ERP analysis for Chapter 5 in Matlab.

This allowed me to analyse my date during lockdown and has greatly contributed to finishing my thesis in time. I also want to thank Frank for answering all my questions about the submission and printing process.

Special mention should also be made of those in Sydney who helped me when I mistakenly decided that we could use printed tests for the pilot. Vince, Judy, Andrea and Leonie kindly helped me sort, assemble and staple over 200 paper tests. During these hours we learnt important lessons, such as how to best set up an assembly line, how after a while, your fingers seem to automatically be able to find the correct page, and most importantly, to never use paper tests again if it can be avoided.

I want to thank those fellow PhD students who have become my good friends over the last three years, as well as my non-PhD friends who supported me throughout this project. You are what makes this work fun. Thanks for all the support, the talks, the laughs, the rants and the general good times: Andrea, Winnie, Pauline, Aida, Hannah, Leonie, Diane, Suzan, Britta, Ana, Frank, Simon, Michiel and others. Special thanks also go to Jidde and Myra for letting me and Izaak live in their apartment while they are in Sydney. It has been the perfect place to start our stay in the Netherlands.

Of course, my academic career would never have existed if it wasn't for the support and encouragement of my parents, Heinz and Angelien, and the general support of my close family: Denise, Chantal, Bart, Babette, Rob, Robert, Karin, Marga and others.

Finally, special thanks go out to my partner, fiancé and now husband Izaak Lea, who moved with me twice for this PhD, once interstate and once intercontinentally, and didn't complain about it once. Thank you for your never-ending positive attitude and support. Due to the pandemic, the final year of work for my PhD took place almost exclusively at home, and your company is what kept me sane.





# CONTENTS

<b>Acknowledgements</b>	<b>VII</b>
<b>Chapter 1</b>	<b>1</b>
General Introduction	
1.1 <i>HISTORY OF PROGRAMMING AND DEVELOPMENT OF PROGRAMMING LANGUAGES</i>	2
1.2 <i>THE COGNITIVE PERSPECTIVE ON PROGRAMMING</i>	4
1.3 <i>THEORETICAL FRAMEWORKS AND COGNITIVE MODELS OF PROGRAMMING</i>	6
1.4 <i>METHODS IN PROGRAMMING EDUCATION RESEARCH</i>	8
1.5 <i>NEUROIMAGING METHODS IN PROGRAMMING LANGUAGE RESEARCH</i>	9
1.6 <i>CONTRIBUTION OF THE THESIS</i>	11
<b>Chapter 2</b>	<b>13</b>
Validating two short versions of the Second Computer Science 1 programming ability test	
2.1 <i>INTRODUCTION</i>	14
2.1.1    Second Computer Science 1 (SCS1) as an assessment tool	15
2.1.2    Need for short versions	16
2.1.3    Need for parallel versions	16
2.1.4    Lack of evidence for external validity	16
2.1.5    Reliability and difficulty	17
2.2 <i>METHODS</i>	19
2.2.1    Participants	19
2.2.2    Materials	20
2.2.3    Procedure	21
2.2.4    Analysis	22
2.3 <i>RESULTS</i>	24
2.3.1    Test difficulty	24
2.3.2    External validity	25
2.3.3    Internal validity: Quality of individual items	26
2.4 <i>DISCUSSION</i>	33
2.4.1    Future research	35
2.4.2    Conclusion	36
<b>Chapter 3</b>	<b>39</b>
Using cognitive skills to predict programming performance following an introductory computing course	
3.1 <i>INTRODUCTION</i>	40
3.2 <i>METHODS</i>	45
3.2.1    Participants	45
3.2.2    Materials	45
3.2.3    Procedure	51
3.3 <i>RESULTS</i>	52

3.3.1	Analysis	52
3.3.2	Pre-processing	52
3.3.3	Regression models	53
3.3.4	Structural equation modelling	57
3.4	<i>DISCUSSION</i>	59
3.4.1	Generalised versus course-related programming performance	60
3.4.2	Lack of predictive ability for pattern recognition	61
3.4.3	Lack of predictive ability for language skills	62
3.4.4	Conclusion	63
<b>Chapter 4</b>		<b>67</b>
Autistic traits and programming learning outcomes in an introductory computing course		
4.1	<i>INTRODUCTION</i>	68
4.2	<i>METHODS</i>	72
4.2.1	Ethics statement	72
4.2.2	Participants	72
4.2.3	Materials	73
4.2.4	Procedure	79
4.2.5	Analysis	80
4.3	<i>RESULTS</i>	81
4.4	<i>DISCUSSION</i>	83
<b>Chapter 5</b>		<b>87</b>
Processing of violations in human and computer languages: An EEG study		
5.1	<i>INTRODUCTION</i>	88
5.1.1	Event-Related Potentials (ERP) and language processing	89
5.1.2	The current study	91
5.2	<i>METHODS</i>	94
5.2.1	Ethics statement	94
5.2.2	Participants	94
5.2.3	Materials	95
5.2.4	Procedure	98
5.2.5	EEG Recording and Data processing	100
5.2.6	Analysis	101
5.3	<i>RESULTS</i>	102
5.3.1	Accuracy data	102
5.3.2	ERP results	102
5.3.3	Summary of ERP results	115
5.4	<i>DISCUSSION</i>	117
5.4.1	Conclusion	120
<b>Chapter 6</b>		<b>123</b>
General Discussion		
6.1	<i>OVERVIEW</i>	124
6.2	<i>ANSWERS TO RESEARCH QUESTIONS</i>	124

6.3	<i>THE ROLE OF NATURAL LANGUAGE SKILLS IN PROGRAMMING</i>	127
6.4	<i>POTENTIAL IMPLICATIONS FOR EDUCATION</i>	129
6.5	<i>SUGGESTIONS FOR FUTURE STUDIES</i>	130
6.5.1	Interactions between course content, test format, and cognitive skills	131
6.5.2	Further disentangling the relationships between cognitive skills and programming	132
6.5.3	Autistic traits and programming	133
6.5.4	How to take advantage of the newly developed ERP presentation method	134
6.6	<i>OVERALL CONCLUSION AND IMPACT</i>	134
	<b>Appendix A</b>	<b>137</b>
	<b>Appendix B</b>	<b>141</b>
	<b>Appendix C</b>	<b>155</b>
	<b>References</b>	<b>167</b>
	<b>Summary</b>	<b>183</b>
	<b>Samenvatting</b>	<b>189</b>
	<b>GRODIL</b>	<b>195</b>



## List of figures

<b>Figure 1.1.</b> PGK-hierarchy model and related skills.	8
<b>Figure 2.1.</b> Median raw scores and distributions per version of the SCS1-S.	24
<b>Figure 2.2.</b> Scatter plot between raw SCS1-S scores and student grades on the course exam.	25
<b>Figure 3.1.</b> Example of a learning item on the vocabulary learning test.	49
<b>Figure 3.2.</b> Example of a learning item on the grammar learning test.	50
<b>Figure 3.3.</b> Partial slopes for each cognitive skill predicting scores on the SCS1 and course grades.	56
<b>Figure 3.4.</b> Structural model with generalised programming skill.	58
<b>Figure 3.5.</b> Structural model with course-related programming skill.	59
<b>Figure 4.1.</b> Example of an item in the vocabulary learning test.	77
<b>Figure 4.2.</b> Example of a learning item from the grammar learning test.	78
<b>Figure 5.1.</b> Examples of stimulus presentation.	100
<b>Figure 5.2.</b> Line plots per region of interest and head plots over time for Dutch stimuli.	106
<b>Figure 5.3.</b> Line plots per region of interest and head plots over time for English stimuli.	109
<b>Figure 5.4.</b> Line plots per region of interest and head plots over time for Java stimuli.	112

## List of tables

<b>Table 2.1.</b> Accuracy and response rates for the two versions of the SCS1-S.	27
<b>Table 2.2.</b> Point biserial correlations of individual items with SCS1-S total score and course grade.	29
<b>Table 2.3.</b> Crosstabulation of item difficulty and discriminability using the scheme from Parker et al. (2016).	30
<b>Table 2.4.</b> Sensitivity of Cronbach's alpha, difficulty and discriminability for each item.	32
<b>Table 3.1.</b> Predictors of score on the SCS1-S programming test and course grade at the end of the semester.	55
<b>Table 4.1.</b> Example items for each subscale of the AQ.	74
<b>Table 4.2.</b> AQ subscales as the predictors of SCS1-Short scores and course grades.	82
<b>Table 4.3.</b> Correlations between AQ score and scores on the cognitive tests.	82
<b>Table 5.1.</b> Examples of the experimental stimuli.	97
<b>Table 5.2.</b> Examples of the fillers.	98
<b>Table 5.3.</b> Summary of ANOVA results by language.	116
<b>Table 5.4.</b> Summary of statistical results comparing grammaticality effects between the languages.	117



