A mixed-methods study into the role of the time variable in the construct of computer-administered C-Tests in three languages

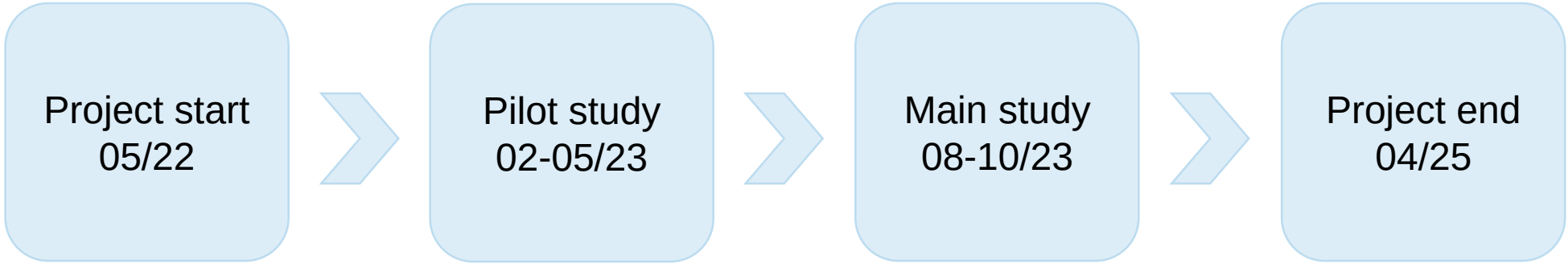
Anastasia Drackert,
Anna Timukova,
Franziska Möller,
Darja Felberg, Ilka Plesse

EuroSLA 32
University of Birmingham

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation
project number 462766474

g.a.s.t.

Project Milestones



Overview

- Introduction to the C-Test and its construct
- Study objectives, design, RQs & methodology
- Operationalization of important constructs
- Results from the pilot study (RQ3)
- Moving forward: The main study

C-Test

2nd half of every 2nd word deleted

4 – 6 short texts

Water

The United Nations Environment Programme has said that water is probably the most serious environmental problem we face apart from climate change . The Earth h as all t he water i t needs t o supply t 6 bil lion people w ho live o f it wi th clean , sa lt water f or drinking . B ut millions o f people m die bec ause of shor tage . This i s because o ur resources a re not loo after properly . The future we are heading for is a world in which many people do not have access to basic , clean water supplies . And instead of working together , we are going to see increasing competition among nations for shared water resources .

20 – 25 gaps

5 mins per text

g.a.s.t.

C-Test

- Integrative measure of **global language proficiency**: „objective, highly reliable and very economical“ (Grotjahn, 2013, p. 181)
- **placement** (e.g. Drackert & Felberg, 2019)
- **screening** before using expensive and time-consuming test batteries (e.g. Eckes, 2014)
- SLA research and studies (Norris, 2018) on educational monitoring (e.g. Harsch & Schröder, 2007)
- quality assurance (e.g. Deutsches Sprachdiplom)

C-Test: Construct

- Modification of the **cloze test** (Raatz & Klein-Braley, 1982)
- Principle of **reduced redundancy**:
  - higher proficiency – less redundancy needed
- **Lower-level** (lexical, morphological, syntactic and orthographic) & **higher-level** skills (awareness of intersentential relationships, metacognitive strategies, global reading skills etc.)
- **„Fluid construct“:** aspects of construct tapped by C-Test depend on **text difficulty & learner proficiency** (Sigott, 2004)

What about the **time**?

C-Test: Construct & Time

Almost all knowledge about C-Test construct based on tests with **generous** time limit

1. Generous time limit of **5 min** per C-Test text
   (e.g. Eckes, 2010; Harsch & Harting, 2015; Porsch & Wilden, 2017)
2. **Reduced** time limit:
   a. reduced & constant for each C-Test text:
      - L1 research on intelligence in the field of psychology (e.g. Raatz, 2002; Wockenfuß, 2008; Wockenfuß & Raatz, 2014)
      - L2 research (e.g. Bisping, 2006; Drackert & Felberg, 2019)
   b. partially variable (e.g. Reichert et al., 2010)
   c. drastically reduced & text-specific = **Speeded C-Test** (e.g. Forthmann et al., 2019; Grotjahn et al., 2010; Heine, 2017; Zimmermann, 2019)

Speeded C-Test

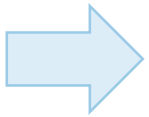
- Grotjahn (2010):
  - Canonical C-Test (5 min per text) measures the **amount of learners‘ declarative and procedural knowledge**
  - Speeded C-Test **additionally** measures the degree of **automaticity** of their skills and the **efficiency of information processing** (cf. p. 285)

Hypotheses:

- S-C-Test would correlate **higher** with measures of **listening** comprehension and **speaking** skills than a canonical C-Test (time pressure)
- S-C-Test would show **lower** correlations with learners‘ **writing** and **reading** skills than a canonical C-Test if measured under generous time conditions (Grotjahn, 2010, p. 289)

Research gaps

- Research focused on:
  - **individual learner groups**, primarily **highly proficient L2 learners** and **native speakers** in L1/L2 German (Grotjahn et al., 2010; Zimmermann, 2019)
  - **correlational analyses** (Fadaeipour & Zohoorian, 2017; Zimmermann, 2019) and **comparisons of test difficulty** (Grotjahn et al., 2010)



unknown changes of the proportion of different aspects of knowledge (declarative and procedural knowledge) & skills (lower- & higher-level processing skills) in the C-Test construct when completion time is drastically reduced

Objective of the study

- Using different methods to gather **various types of evidence** to answer a range of research questions to specifically investigate the **role of the time variable** in the **C-Test construct** in a comprehensive way to allow for a higher degree of generalizability of the results for:

  - **computer-administered** C-Tests
  - learners of **different levels of proficiency** (from beginners to advanced)
  - **several languages** (English, German, Russian)

Research questions & methods

RQs	Methods
1. How does the time variable influence the **reliability** of computerised C-Tests?	IRT reliability coefficients
2. How does the time variable influence **learners' scores** depending on their **proficiency level**?	MANCOVA analysis
3. Which components of L2 proficiency (**declarative, procedural knowledge** and **automaticity**) are better predictors of differently timed C-Tests?	Linear regression analysis; SEM
4. How does the time variable influence the **correlations** between a C-Test and an integrated measure of **oral proficiency?**	Correlation (with OEIT)
5. How does the time variable influence the **strategies** deployed by learners?	Process-oriented video-based analysis

Study design

- 2 C-Tests: a canonical and a speeded version
    - 5 texts with 20 gaps in each version
- Oral Elicited Imitation Test
- 7 tests of declarative and procedural knowledge
- test of typing skills (https://10fastfingers.com/)
- background questionnaire

Platforms:

- g.a.s.t.-*Moodle*
- *testable*

Sample

- academic L2 learners
- different proficiency levels
- ENG *N* = 34
- GER *N* = 21
- age *M* = 25

RQ 3:
Which components of L2 proficiency (**declarative, procedural knowledge** and **automaticity**) are better predictors of differently timed C-Tests?

Conceptualising declarative, procedural and automatised knowledge

- **ACT (adaptive control of thought) theory** (Anderson, 1983; Anderson & Lebriere, 1998)
  - general theory of skill learning; transition from DK (knowledge *that;* stored as chunks) to PK (knowledge *how*; "production rules") and automatic execution

- **Skill Acquisition Theory** (DeKeyser 1997; 2014)
  - focus on L2 in language instruction settings; fixed chronological sequence (declarative stage > proceduralisation > automaticity)

- **Neurolinguistic Theory of Bilingualism** (Paradis, 2009)
  - declarative memory conscious (facts and events), while procedural memory unavailable to conscious recall (perceptual, motor or cognitive skills) in different brain regions

Conceptualising declarative, procedural and automatised knowledge

- **Declarative/Procedural Model** (Ullman, 2020; Morgan-Short & Ullman, 2022)
  - Neurobiological model of language learning, knowledge and use
  - **declarative** memory (DM): explicit & implicit knowledge; lexicon (open-class **content words)**, sound-meaning mappings; **irregular morphological** forms; **idiosyncratic** or individual **chunks; generalized** analogies and **explicit** rules
  - **procedural** memory (PM): **implicit** knowledge only; cognitive and (perceptuo-)motor skills, categories, habits; (rule-governed) **grammar** (phonology, morphology and morphosyntax); **predictable** sequences, (**real-time**) combination of elements made possible by prediction of downstream elements
  - knowledge can be **automatized** in **both** memory systems

Conceptualising declarative, procedural and automatised knowledge

- **Declarative/Procedural Model** (Ullman, 2020; Morgan-Short & Ullman, 2022)
  - consequences for **operationalisation**:
  - type & form of **declarative** and **procedural** knowledge „are often quite different, even while this knowledge underlies the **same or similar outcomes**“ (Ullman, 2016, p. 957)
  - **manipulating context/task parameters** can affect which system is relied on more (e.g. reduced attention to stimuli & complexity of rules & patterns PM)

Conceptualising declarative, procedural and automatised knowledge

- **Operationalisation**:
  - **DK**: **consciously** accessible linguistic knowledge (**no time pressure**; attention to the stimuli) stored as chunks, i.e. vocabulary units, explicit grammar & orthography rules
  - **PK**: unconscious (**implicit**) largely automatised knowledge and psychomotor skills necessary for real-time **(online) processing** and production of rule-governed morphological and syntactic sequences
  - **Automaticity**: processing speed and accuracy

RQ 3:
Which components of L2 proficiency (**declarative, procedural knowledge** and **automaticity**) are better predictors of differently timed C-Tests?

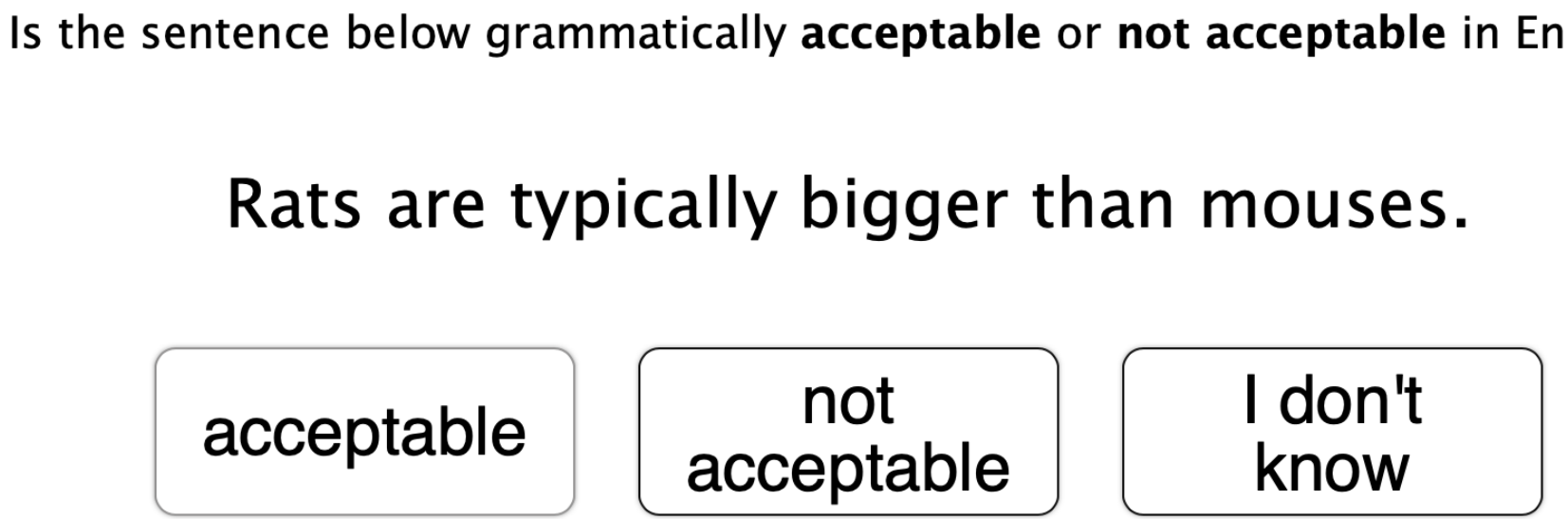
- **Hypothesis**: Performance on the Canonical C-Test can be better predicted by measures of declarative and procedural knowledge, whereas performance on the Speeded C-Test can be better predicted by measures of (procedural knowledge and) automaticity.

- **Planned Method**: Linear regression analysis, SEM

Measures of declarative and procedural knowledge

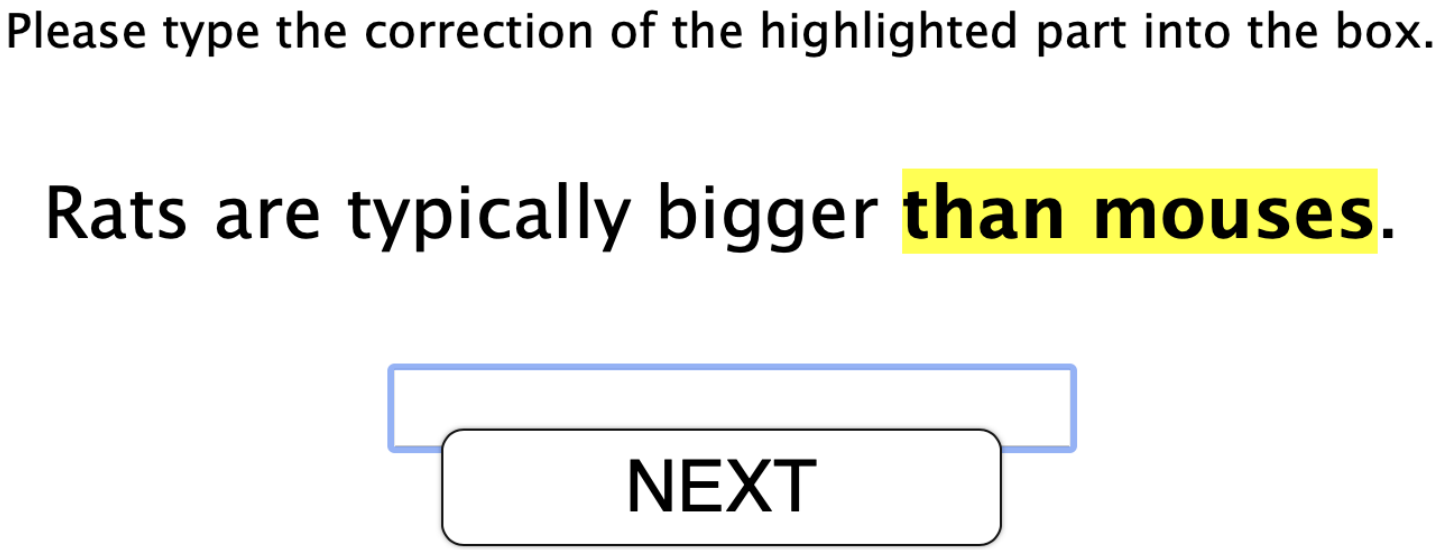
Test	Format	Construct	Source/Author
Vocabulary Size Test (VST)	Match words to definitions (*untimed*)	Declarative (receptive) knowledge of vocabulary (breadth of vocabulary)	Institut für Testforschung und Testentwicklung e.V. Leipzig (Nation, 1990)
Grammatical Acceptability Judgment Test (GAJT)	Decide whether sentences are grammatically acceptable or not (*untimed*)	Declarative (receptive) knowledge of grammar	DeKeyser (2000) & Lu (2010)
Grammar Correction Task (GCT)	Correct highlighted parts of sentences (*untimed*)	*Declarative(?)* (productive) knowledge of grammar	ungrammatical sentences from GAJT
Orthographic Awareness Task (OAT)	Decide whether pseudowords are possible in the target language (*untimed*)	Declarative (abstract) knowledge of orthography (legal letter combinations of a writing system)	Drackert et al. (project); concept by Möller (van der Leij, Bekebrede & Kotterink 2010; König, Calude & Coxhead 2020)
Test	**Format**	**Construct**	**Source/Author**
Orthographic Choice Task (OCT)	Decide whether words are spelled correctly or not (*timed*)	*Procedural(?)* (word-specific) knowledge of orthography	Drackert et al. (based on Olson et al., 1994)
Modified Self-Paced Reading Test (SPRT)	Read sentences part by part; answer questions about their content (distractors) and grammaticality (items) (*timed*)	Procedural (receptive) knowledge of grammar	versions of sentences used in GAJT (targeting same phenomena) (Marsden et al., 2017)
Written Elicited Imitation Test (WEIT)	Reconstruct written stimuli in writing (*timed*)	Procedural integrated linguistic knowledge & skills	Drackert et al. (project); concept by Timukova

Measures of declarative and procedural knowledge: GAJT



- 62 items
- pairs of grammatical / ungrammatical sentences
- different grammatical phenomena
- randomized order of presentation

Measures of declarative and procedural knowledge: GCT



- 32 items
- ungrammatical sentences from GAJT
- randomized order of presentation

Measure of automaticity

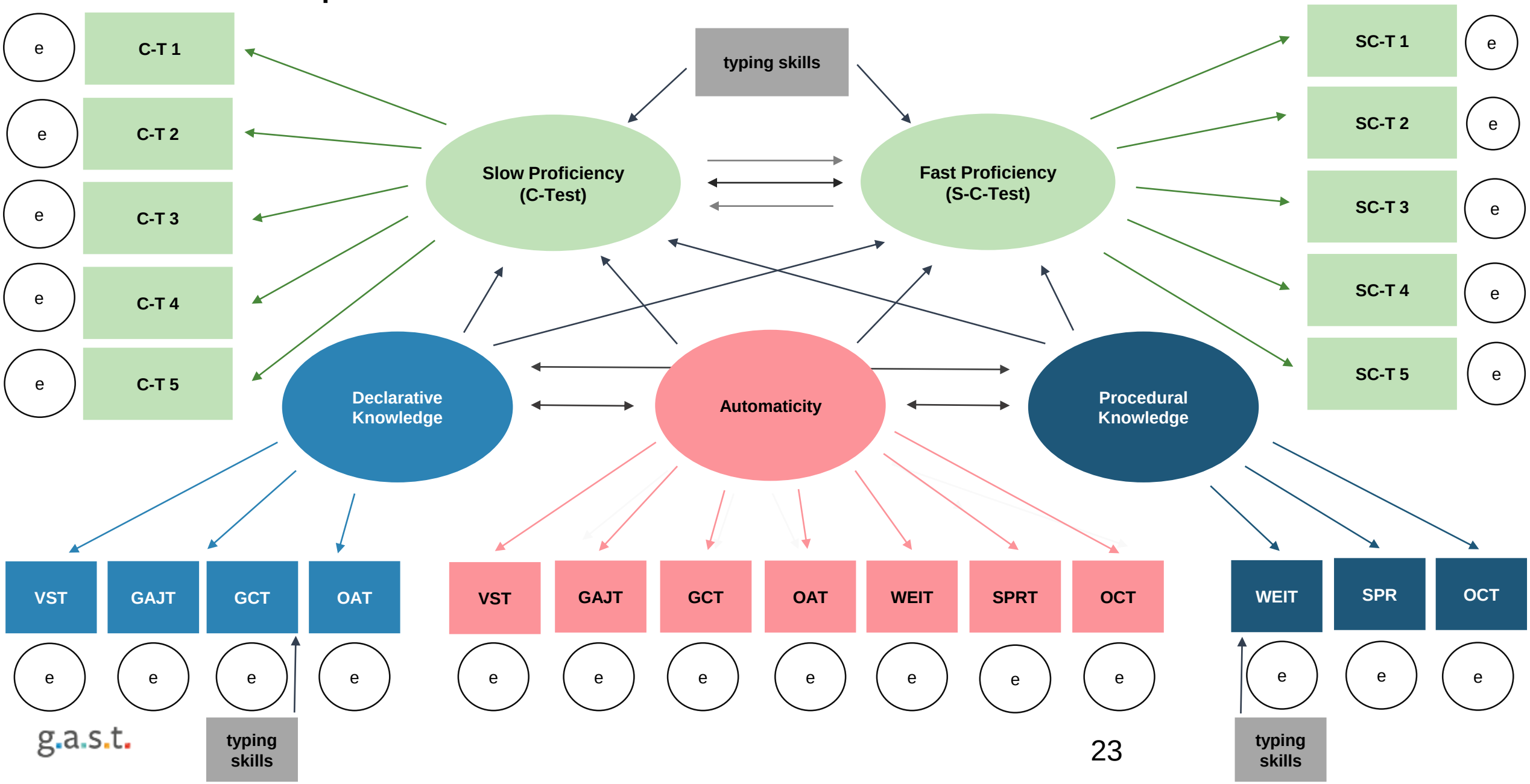
- processing speed and accuracy
- scores and reaction times for correctly solved items

total score on a test / mean reaction time for correctly solved items

Example:

ID	GAJT_score	GAJT_RT	GAJT_Automaticity
pe0103_03	62	2693	.023
pe0103_01	62	4648	.013
pe0103_01	52	13767	.004
pe2402_11	33	7310	.005

Structural Equation Model



g.a.s.t.

Results from the pilot study (RQ 3)

Correlations with scores on different instruments

Correlations C-Test & Speeded C-Test with scores on different instruments ENG

		VST	GAJT	GCT	OAT	OCT	SPR	WEIT
C-Test	Spearman‘s rho	0.832	0.851	0.834	0.769	0.769	0.769	0.856
	p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
S-C-Test	Spearman‘s rho	0.733	0.876	0.866	0.746	0.745	0.756	0.913
	p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Correlations C-Test & Speeded C-Test with scores on different instruments GER

		VST	GAJT	GCT	OAT	OCT	SPR	WEIT
C-Test	Spearman‘s rho	0.627	0.809	0.915	0.768	0.828	0.862	0.962
	p	0.003	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
S-C-Test	Spearman‘s rho	0.608	0.791	0.932	0.739	0.831	0.857	0.979
	p	0.004	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Correlations with automaticity on different instruments

Correlations C-Test & Speeded C-Test with automaticity on different instruments ENG

		VST	GAJT	GCT	OAT	OCT	SPR	WEIT
C-Test	Spearman‘s rho	0.833	0.423	0.802	0.121	0.513	0.708	0.845
	p	<0.001	0.014	<0.001	0.501	0.003	<0.001	<0.001
S-C-Test	Spearman‘s rho	0.812	0.557	0.849	0.163	0.528	0.747	0.906
	p	<0.001	<0.001	<0.001	0.364	0.002	<0.001	<0.001

Correlations C-Test & Speeded C-Test with automaticity on different instruments GER

		VST	GAJT	GCT	OAT	OCT	SPR	WEIT
C-Test	Spearman‘s rho	0.844	0.500	0.845	-0.264	0.374	0.431	0.874
	p	<0.001	0.021	<0.001	0.247	0.095	0.065	<0.001
S-C-Test	Spearman‘s rho	0.853	0.520	0.859	-0.252	0.384	0.477	0.917
	p	<0.001	0.016	<0.001	0.270	0.086	0.039	<0.001

Summary

mixed results of correlation analyses (intended methods not possible with the pilot sample):

 higher correlations for **S-C-Tests** with **automaticity** measures

- **scores on declarative knowledge:** better predictor for performance on **Canonical C-Test**
- **scores procedural knowledge:** no clear tendency (yet)
- **automatised knowledge:** better predictor for performance on **Speeded C-Test**

Moving forward: The main study

- **Theoretical issues**:
  - Time limits on Speeded C-Tests
  - Construct of some instruments (GCT & OCT)
  - Automaticity measure?

- **Practical issues**:
  - Modifications of instruments based on item analyses
  - Acquisition of participants (your support is most welcome!)

August – November 2023

online data collection

participants: N = **540** (**180** per language – three proficiency levels a **60** participants)

Thank you!

Vielen Dank!

Спасибо!

Contact: moeller@gast.de

More info about our study:

ENG

GER

g.a.s.t.