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

g.a.s.t.

Overview

- Object of the study: C-Test & its construct

- Study objectives, design, RQs & methodology

- Results of the pilot study (ENG)

- Towards the main study

C-Test

20 – 25 gaps

2nd half of every 2nd word deleted

4 – 6 short texts

5 mins per text

Media, Please Leave Us Alone

Mainstream media continually reports on the expanding waistlines of Canadians. At t[he] ✔ sam[e] me ✔ that doc[uments] ✖ tell u[s] ✔ about overv[iew] ✖ , heart dis[aters] ✖ and can[cer] ✔ , magazines a[nd] ✔ movies cont[inue] ✔ to fea[ ] ✖ skinny mod[ ] ✖ and actr[ ] ✖ . When yo[ung] ✔ girls g[o] ✔ through pub[lic] ✖ they ga[in] ✔ weight aro[und] ✔ their hi[ ] ✖ but th[ere] ✖ is a si[ ] ✖ of beco[ming] ✔ a woman. T[he] ✔ media tr[ies] ✔ to swi[p] ✖ the nat[ional] ✖ process into a constant fight against it: we get bombarded with images of the perfect body everywhere we turn....

g.a.s.t.

3

C-Test: Construct

- Modification of the cloze test (Raatz & Klein-Braley, 1982)

- Principle of reduced redundancy: higher proficiency – less redundancy needed

- Integrative measure of global language proficiency: „objective, highly reliable and very economical" (Grotjahn, 2013, p. 181)

- Lower-level skills (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**?
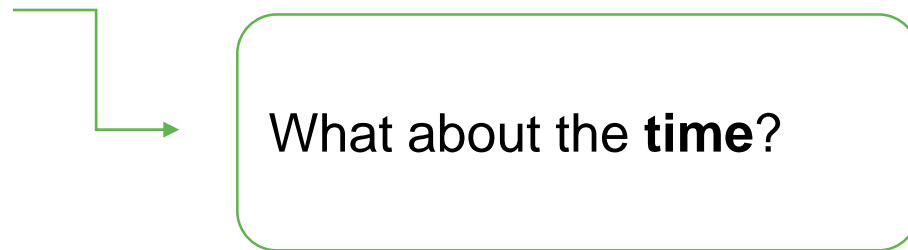
C-Test: Time & the construct

1. Generous time limit of 5 min per C-Test test (e.g. Eckes, 2010; Harsch & Harting, 2015; Porsch & Wilden, 2017)

1. 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; Gortjahn et al., 2010; Heine, 2017; Zimmermann, 2019)

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

Speeded C-Test: Hypotheses about the construct

- 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 (p. 289)

g.a.s.t.

Speeded C-Test: Previous research on time variable

1:05 – 1:55
S-C-Test

2:30

3:00

4:00

5:00
C-Test
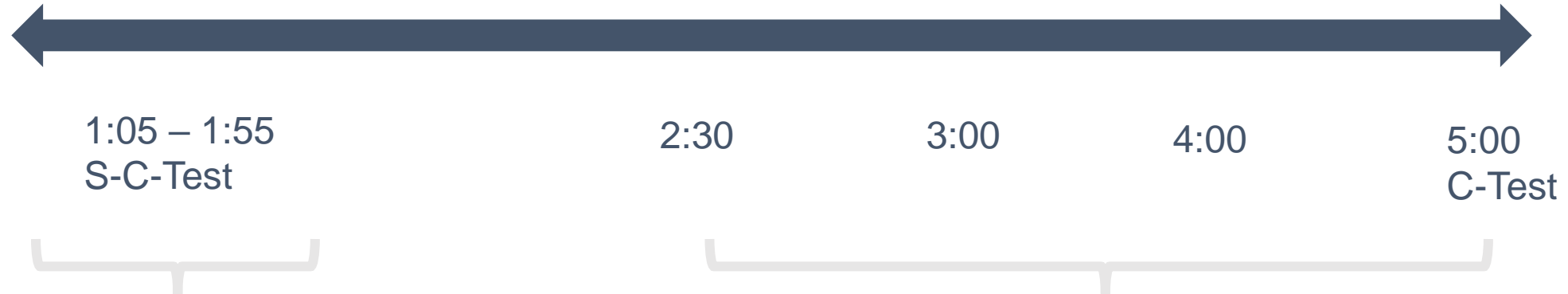
Zimmermann, 2019:

B2 German learners
higher correlations for S-C-Test with tests of **speaking** (dialogical) and **listening comprehension**

Fadaeipour & Zohoorian, 2017:

mixed-level English learners
S-C-Test better predictor of **reading comprehension** (RC Test)
only **two** C-Test **texts** per administration; time limit **2.5 min** for "speeded" text

g.a.s.t.

Research gaps

- **paper-based** C-Tests

- time-reduced **C-Tests as instruments** (Reichert et al., 2010) and/or **not canonical** C-Tests (Raatz, 2002: 1 text 100 gaps instead of 4x25)

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

Objective of the study

- Using different methods to gather various types 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:
  - learners of different levels of proficiency (from beginners to advanced)
  - several languages (English, German, Russian)
  - computer-administered C-Tests

g.a.s.t.

Motivation & relevance

- Elusive construct (**different constructs**?)

- Relatively **easy to develop & use** in local contexts (e.g. to be used for **placement**)

- **Modifications** can be made to construction principles, scoring & time to adjust to the **purpose, target group** and **target language** (e.g. Drackert & Timukova 2020: heritage speakers of RUS)

- Stake holders need **quick(er) tests** (no justification for 5 minutes; piloting new texts; ChatGPT more difficult to use under time pressure?)

Research questions & methods

RQ	Method(s)
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 (Kerschhofer-Puhalo, Lalouschek & Mayer, 2018)

g.a.s.t.

Study design

- Two C-Tests: Version A (speeded) & Version B; 5 texts with 20 gaps in each version

- Oral Elicited Imitation Test

- Battery of 7 tests of declarative and procedural knowledge (for RQ3)

- Test of typing skills (https://10fastfingers.com/)

- Background questionnaire

Platforms:

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

- *testable*

Setting **time limit per text** for speeded C-Test:

- pilot with native speakers;
- average processing times per text + 20% -> rounded up;
- from **1:40 min** to **2:50 min** for all languages (still open for discussion)

g.a.s.t.

13

Pilot results for RQ1, RQ2 & RQ4

- ENG sample:
- $N = 34$,
- age $M = 25{,}92$,
- students,
- different L1s

g.a.s.t.

RQ1: How does the time variable influence the reliability of computerised C-Tests?

- **Hypothesis**: The reliability of the C-Test is expected to be influenced by the time factor and moderated by learners' L2 proficiency.

- **Method**: Cronbach's alpha *(main study: IRT reliability coefficients)*

Reliability C-Test & Speeded C-Test (ENG)

	N	Cronbach's alpha	*N* of items
C-Test	34	.949	5
Speeded C-Test	34	.951	5

g.a.s.t.

RQ2: How does the time variable influence learners' scores depending on their proficiency level?

- **Hypothesis**: Learners' scores are expected to increase with additional time. The amount of gain in the scores will depend on learners' level of proficiency.

- **Method**: Wilcoxon Signed Rank Test *(main study: MANCOVA analysis)*

Total scores C-Test & Speeded C-Test (ENG)

	N	*M*	*SD*	*Min.*	*Max.*
C-Test	34	70.74	20.60	15	94
Speeded C-Test	34	68.44	20.63	15	93

$z = -1.996$, $p = 0.046*$

g.a.s.t.

RQ4: How does the time variable influence the correlations between a C-Test and an integrated measure of oral proficiency?

- **Hypothesis**: Completion of a C-Test under time constraints will require learners to rely largely on their automatized language knowledge, so the less time is available for a C-Test, the higher it will correlate with OEIT. The results are expected to be influenced by learners' L2 proficiency.

- **Method**: correlations with OEIT

Correlation C-Test & Speeded C-Test with OEIT ENG

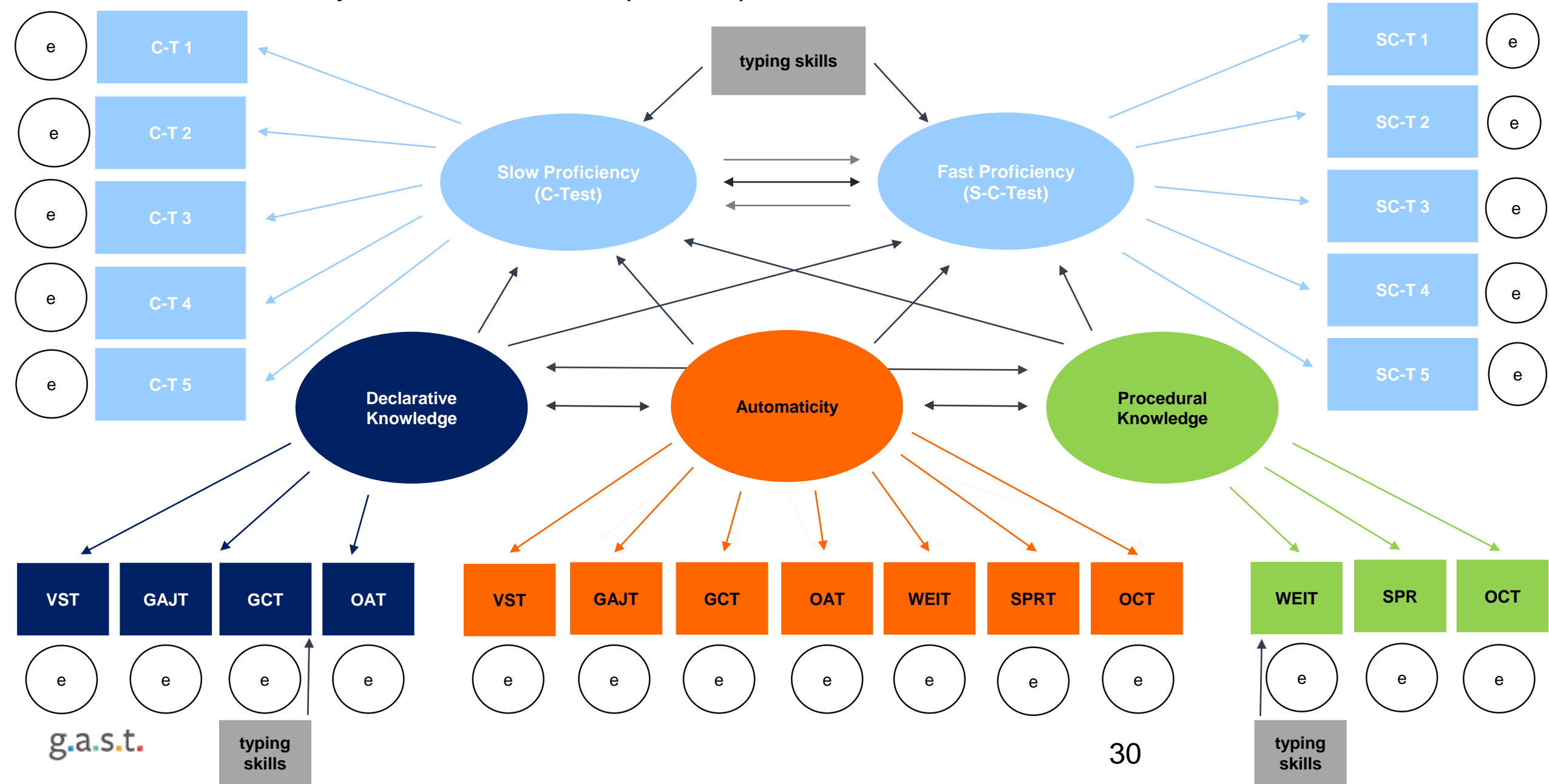|	*N*	**Spearman's rho**	*z*	*r²*
C-Test	34	.740	<.001	.548
Speeded C-Test	34	.827	<.001	.684

RQ 3: How does the time variable influence the role of different components in the construct of a C-Test?

- **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

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

Instruments: Declarative and procedural knowledge (RQ 3)

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)
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

g.a.s.t.

29

Structural Equation Model (RQ 3)

RQ 3: How does the time variable influence the role of different components in the construct of a C-Test?

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 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

g.a.s.t.

Summary of the pilot results

RQ1: reliability of C-Test & S-C-Test ENG almost the same (influence of L2 proficiency cannot be determined yet)

RQ2: learners' scores increase significantly with additional time (influence of L2 proficiency cannot be determined yet)

RQ3: mixed results of correlation analyses (intended methods not possible with the pilot sample): higher correlations for S-C-Tests with automaticity measures

RQ4: S-C-Test correlates higher with OEIT in ENG (influence of L2 proficiency cannot be determined yet)

RQ5: not discussed

Towards the main study

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

- **Practical issues**:
    - Modify instruments based on item analyses
    - Plan online administration
    - Recruite participants (your support is most welcome ☺!)

July – November 2023

Online data collection

participants: n = **540** (**180**

per language – three

proficiency levels a **60**

participants)

Towards the main study

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

- **Practical issues**:
  - Modify instruments based on item analyses
  - Plan online administration
  - Recruite participants (your support is most welcome ☺!)

July – November 2023

Online data collection participants: n = **540** (**180** per language – three proficiency levels a **60** participants)

g.a.s.t.

Thank you!
Vielen Dank!
Спасибо!

anna.timukova@gast.de

ENG

GER

g.a.s.t.