Writing-to-Learn Research

Veerle Baaijen
Statistic & Methodology
Seminar
16 mei 2007

Content

1. Writing Process Research: a short introduction

2. How to measure the writing processes?

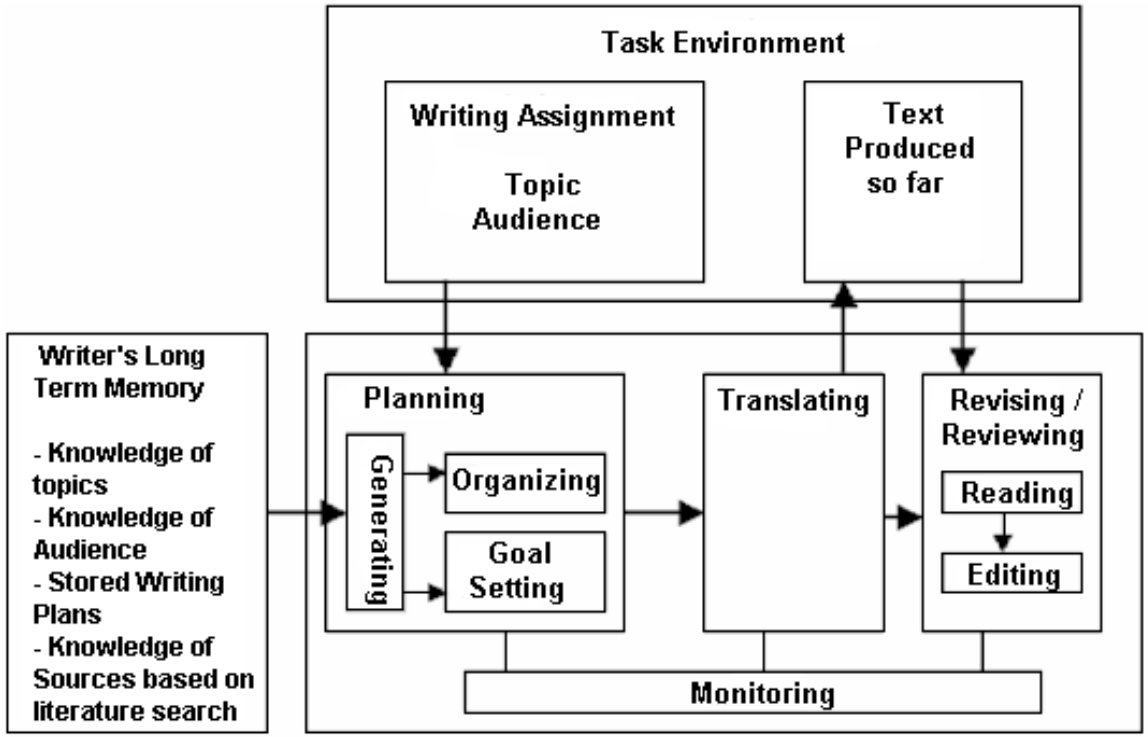
3. Observations of writing processes and knowledge change: an evaluation of thinking aloud protocol analysis

4. Non- parametric tests: Mann-Whitney & Kruskal Wallis test

1. Writing Process Reseach

- Writing- to- learn:
  - Over several decades writing has been considered as a powerful learning task
  - Writers develop new ideas as they write
  - How can writing contribute to learning?

Flower and Hayes



Expert vs. Novice writers

- Novice writers → knowledge telling
- Expert writers → knowledge transforming
- Process of writing →knowledge change → problem solving processes.

Bereiter and Scardamalia

- Knowledge Telling
  - Children and novice writers
  - ‘Think-say’ method of composing
  - Writers pay a lot of attention to the surface features of the text
  - Planning & Revising not important
- Knowledge Transforming
  - Expert adult writers
  - Problem-solving activity
  - Communicative goals with respect to the reader
  - Planning & Revising important

A definition of writing process research

The study of “the otherwise hidden traces [of text production], the concealed ‘seams’, … which cannot be observed from analysis of the written product alone” (Spelman Miller,2000)

2. How to measure the writing processes?

- Different methods:
  - Verbal and introspective reporting:

    “gathering data by asking individuals to vocalize what is going on through their minds as they are solving a problem or performing a task” (Gass & Mackey, 200, p.13)

  -Observation: video-tapes of writing behavior

  -Keystroke-logging: registration of keystroke movement of writer

Thesis Research

- Combination between think aloud protocols and keystroke-logging
- To which extend do think aloud protocols provide information about the mental processes during writing?

Design Experiment

- 24 secondary education students (15-16 years old)
- 13 male – 11 female
- Two assignments
- Two groups which were stratified by gender (male-female) and level of ability (level 1,2 & 3)

Counterbalanced Design

Research Design: Sex * Assignment

Count

		Opdracht		
		HOD 1	HOD 2	Total
Geslacht	man	7	6	13
	vrouw	5	6	11
Total		12	12	24

Research Design: Level of ability * assignment

Count

		Opdracht		
		HOD 1	HOD 2	Total
Niveau	Hoog	4	4	8
	Middel	4	4	8
	Laag	4	4	8
Total		12	12	24

Writing Assignments

	Writing Assignment 1	Writing Assignment 2
Audience	Junior students	Junior students
Topic	Harmful effects of cfc's (chlorofluorocarbons)	Opposite effects of a volcano eruption and an oil burn on temperature.
Goal	Informative	Informative
Concepts to be used	Ozone layer, filter, ultraviolet radiation	Atmosphere, greenhouse effect, carbon dioxide

Example logfile

- 202.6 I 17 2.7 om de aarde heen
- 9.9 I 34 9.6 bevind zich een atmosfeer, dit iss
- 3.1 D -1
- 13.0 I 50 12.6 een luchtlaag die de aarde beschermt tegen het zo
- 8.2 I 22 8.0 nlicht en ruimtepuin.
- 15.3 I 5 1.4 koold
- 0.8 D -1
- 10.8 I 33 10.4 stofdioxide is een stof net zoald
- 0.3 D -1
- 2.5 I 11 4.3 s zuurstof
- 2.7 D -1
- 7.5 I 28 7.2 , die belangrijk is voor de
- 7.8 I 15 3.8 arde en ons leb
- 0.6 D -1
- 2.6 I 5 2.1 ven.

- Information about pausing
- Formulating activities
- Revising activities
  - **Keystrokes:**

  - I= Insertion
  - D=Deletion
  - M=Movement

Example think aloud protocol

Lezen opdracht (hardop)	
Hardopdenken	nou dat moet lukken ehm
Hardopdenkend formuleren	ehm nou eh om de aarde heen eh bevindt zich een atmosfeer dit is een luchtlaag eh die de aarde beschermt tegen het zonlicht en ruimtepuin
Hardopdenken	eh ehm ja het broeikaseffect eh dan ga ik eerst over koolstofdioxide denk ik maar
Hardopdenkend formuleren	eh koolstofdioxide is een stof net zoals zuurstof die belangrijk is voor de eh ja aarde en ons leven eh
Hardopdenken	eh met die vulkaanuitbarsting eh komt er veel veel eh komt er veel koolstofdioxide vrij
Hardopdenkend formuleren	uh het probleem met onze atmosfeer is dat
	eh mensen ja ehm
Herlezen tekst (in stilte)	

47.0						lezen opdracht (hardop)	20
3.0					nou dat moet lukken ehm	hardopdenken	1
11.0					ehm nou eh om de aarde heen	hardopdenken	9
	I	17	2.7	om de aarde heen	eh bevindt zich een atmosfeer dit is een luchtlaag eh die de aarde beschermt tegen het zonlicht en ruimtepuin	hardopdenkend formuleren	10
9.9	I	34	9.6	bevind zich een atmosfeer, dit iss			
3.1	D	-1		[Delete 's']			
13.0	I	50	12.6	een luchtlaag die de aarde beschermt tegen het zo			
8.2	I	22	8.0	nlicht en ruimtepuin.			
15.3					eh ehm ja het broeikaseffect eh dan ga ik eerst over koolstofdioxide denk ik maar	hardopdenken	4
	I	5	1.4	koold	eh koolstofdioxide is een stof net zoals zuurstof die belangrijk is voor de eh ja aarde en ons leven eh	hardopdenkend formuleren	10
0.8	D	-1		[Delete 'd']			
10.8	I	33	10.4	stofdioxide is een stof net zoald			
0.3	D	-1		[Delete 'd']			
2.5	I	11	4.3	s zuurstof			
2.7	D	-1		[Delete spatie]			
7.5	I	28	7.2	, die belangrijk is voor de			
7.8	I	15	3.8	arde en ons leb			
0.6	D	-1		[Delete 'b']			
2.6	I	5	2.1	ven.			
26.7					eh met die vulkaanuitbarsting eh komt er veel veel eh komt er veel koolstofdioxide vrij	hardopdenken	2
	I	28	7.7	het problem met onze atmos v	uh het probleem met onze atmosfeer is dat	hardopdenkend formuleren	10
6.6	D	-1		[Delete 'v']			
2.8	I	12	2.4	feer is dat			

Quantitative analysis: coding scheme

Code	Categorie
Planning	
1	Task Orientation
2	Generating Ideas
3	Setting Procedural Goals
4	Setting Content Goals
5	Setting Rhetorical Goals
7	Setting Structural Goals
Translating	
8	Taking Notes
9	Memorial Pre Text
10	Formulating
Reviewing	
11	Reviewing memorial Pre-text
12	Reviewing surface features
13	Reviewing Procedural Goals
14	Reviewing Content Goals
15	Reviewing Rhetorical Goals
17	Reviewing Structural Goals
Reading	
19	Reading sources
20	(re) Reading the task
21	(re) Reading one's own text (partly)
23	(re) Reading one's own text (whole)

- All thought units were coded
- 4 Main categories
  -Planning
  -Translating
  -Reviewing
  -Reading
- 19 subcatogories

- Example: Idea generation (1) After reading the assignment: "hm, why would the temperature drop as a lot of heat will be set free. (2) Let's see. May be it has to do with the damp"

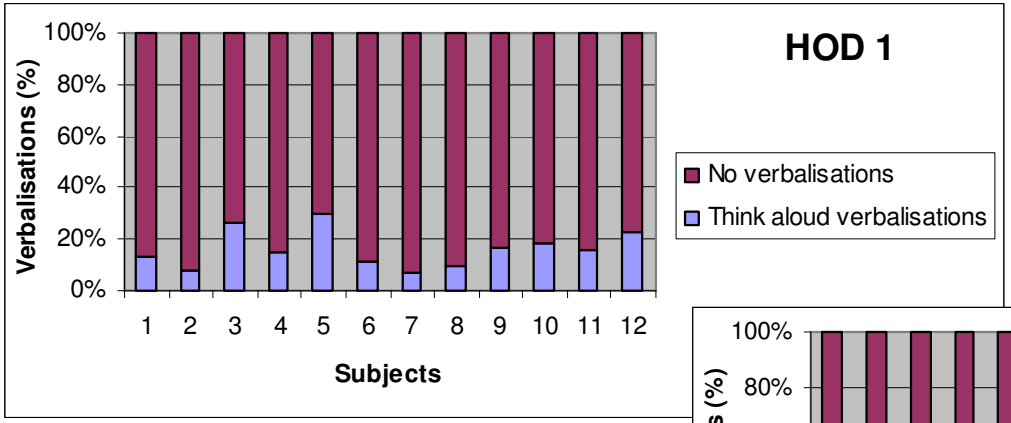
3. An evaluation of think aloud protocol analysis

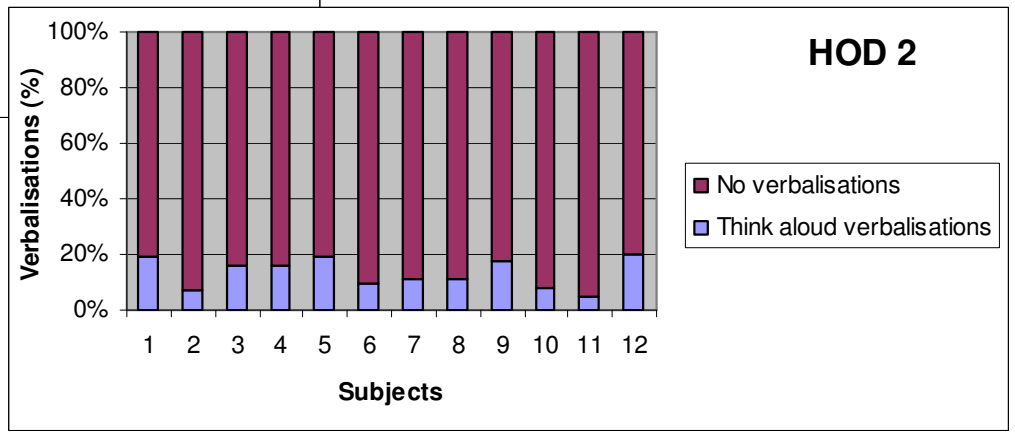
1 What part of the pauses are filled with verbalisations?

2. With which mental processes are these pauses filled?

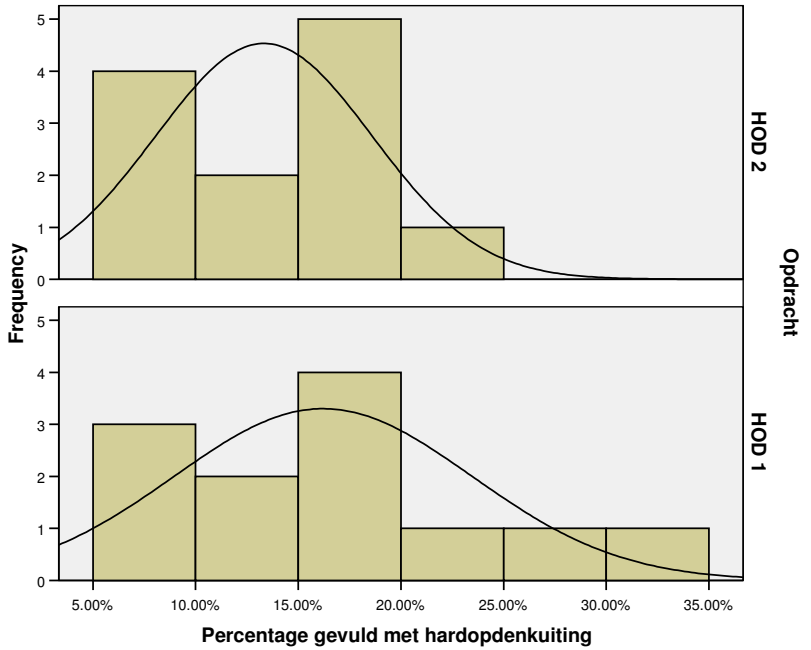
3. Do sex, assignment and level of ability matter?

1. Verbalisations





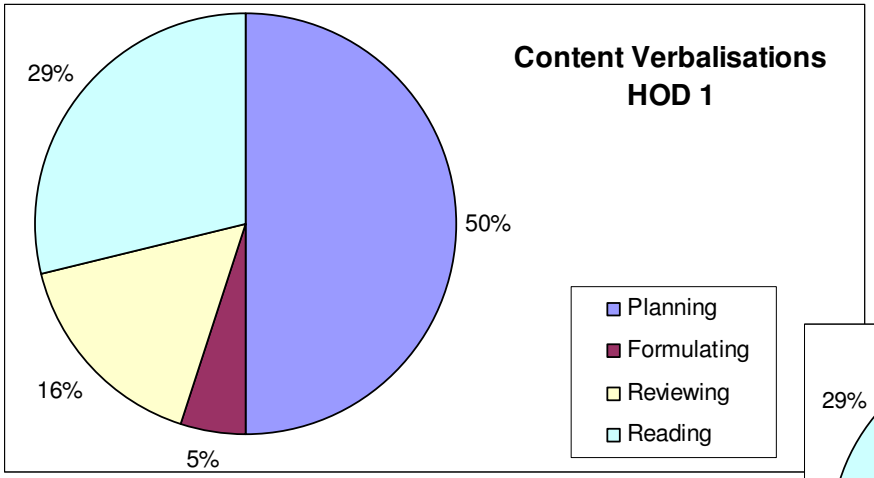
Descriptives and Distribution

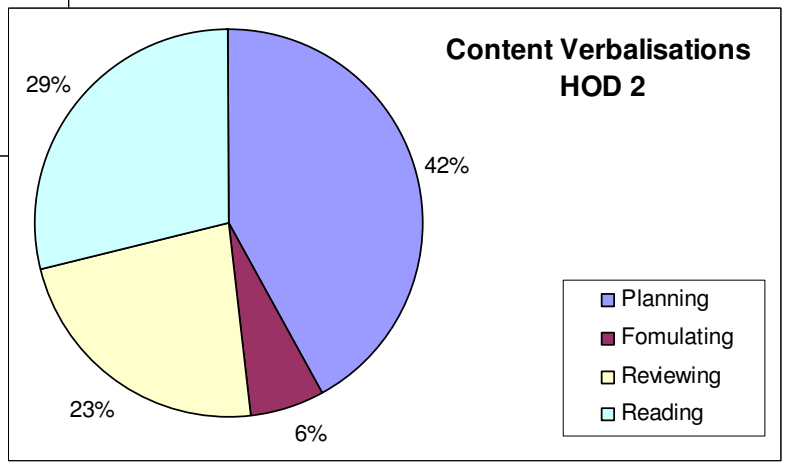


Percentage gevuld met hardopdenkuiting

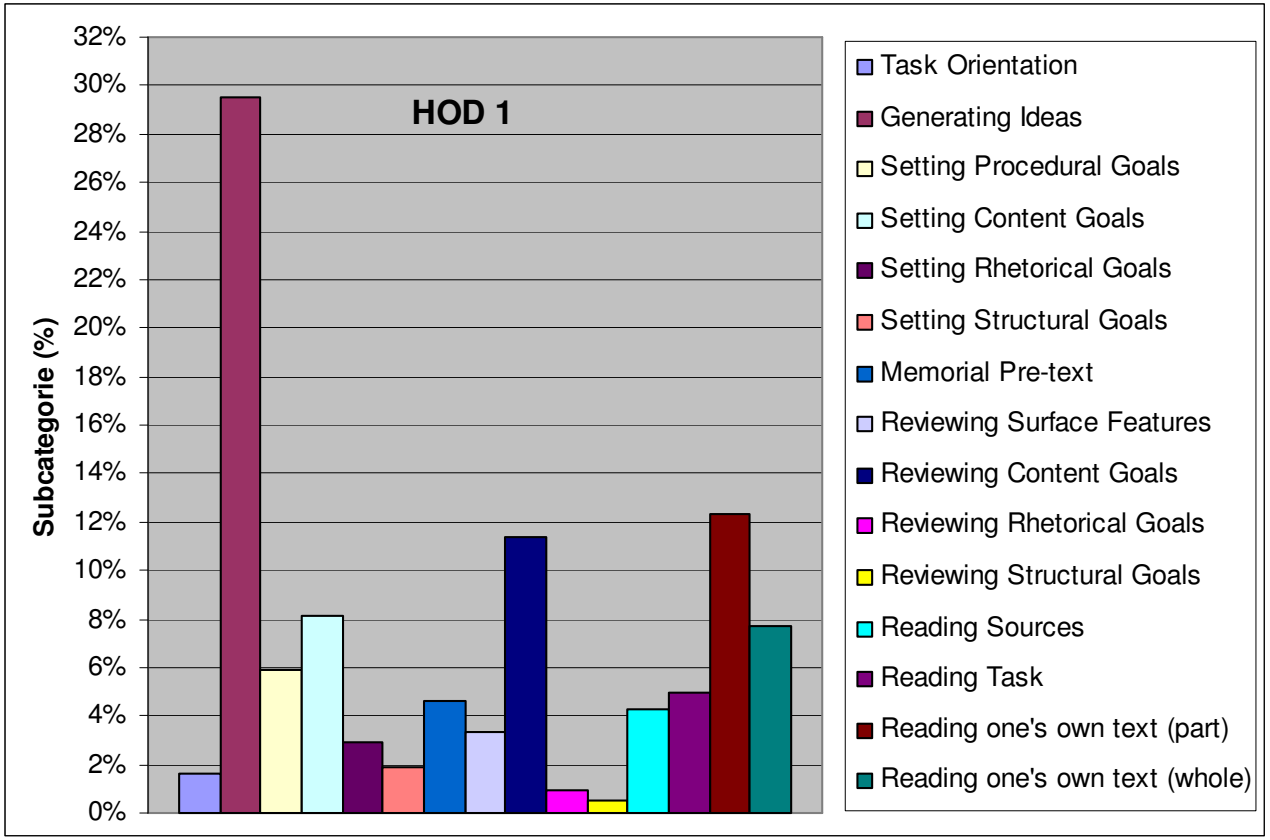
Opdracht	Mean	N	Std. Deviation	Minimum	Maximum	Range	Median
HOD 1	16.2098%	12	7.24992%	6.98%	30.05%	23.08%	15.6551%
HOD 2	13.3274%	12	5.27968%	5.08%	20.00%	14.92%	13.5802%
Total	14.7686%	24	6.37471%	5.08%	30.05%	24.97%	15.5666%

2. Content verbalisations

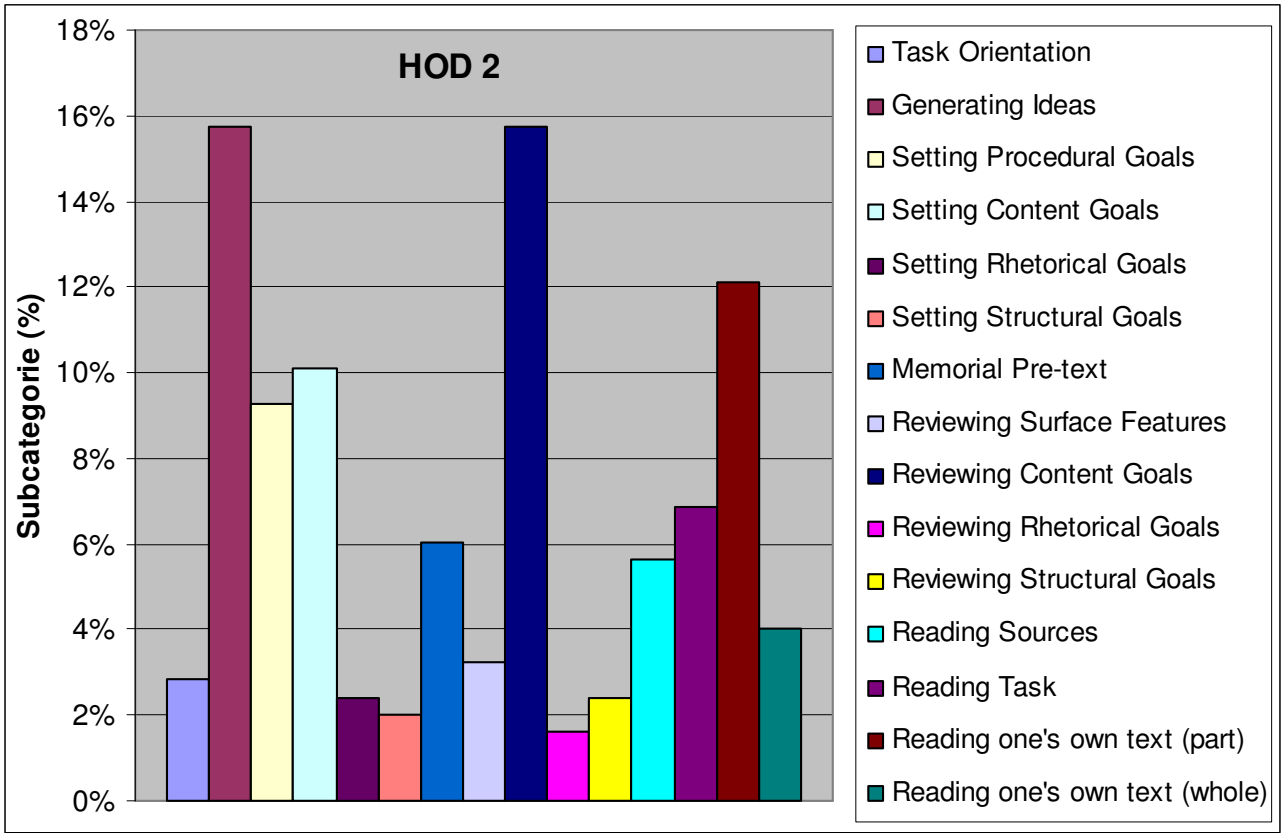




Codes


HOD 1
Subcategorie (%)
32%
30%
28%
26%
24%
22%
20%
18%
16%
14%
12%
10%
8%
6%
4%
2%
0%
Task Orientation
Generating Ideas
Setting Procedural Goals
Setting Content Goals
Setting Rhetorical Goals
Setting Structural Goals
Memorial Pre-text
Reviewing Surface Features
Reviewing Content Goals
Reviewing Rhetorical Goals
Reviewing Structural Goals
Reading Sources
Reading Task
Reading one's own text (part)
Reading one's own text (whole)

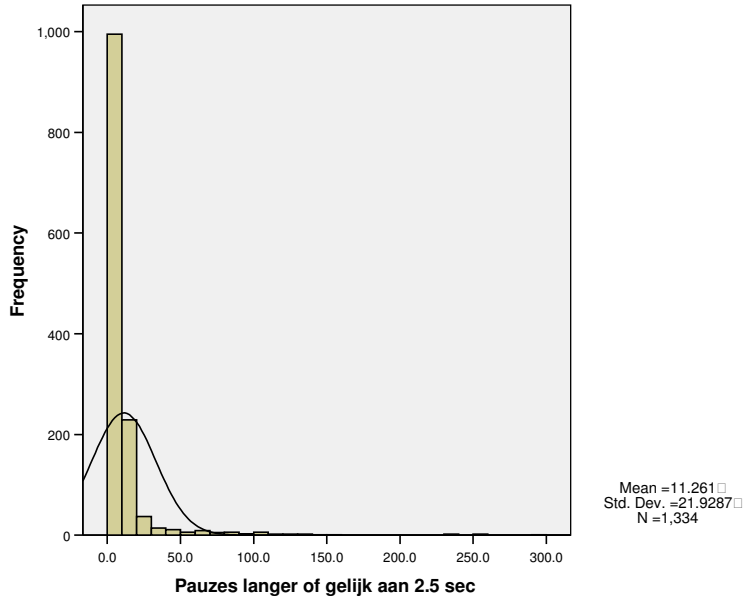
Codes

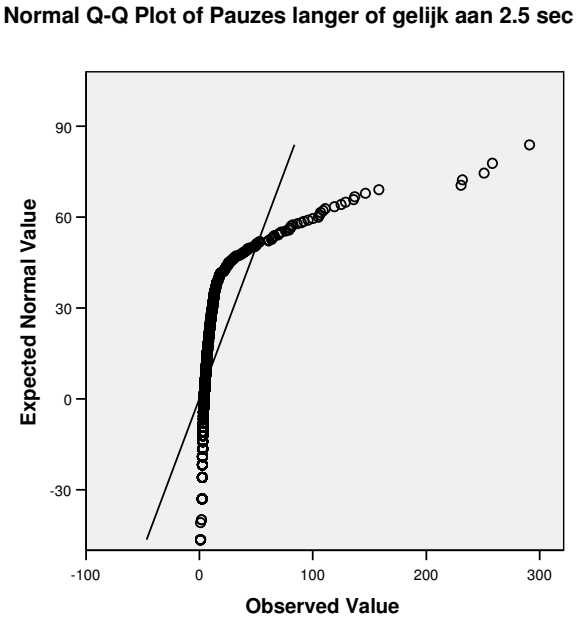

HOD 2
Subcategorie (%)
18%
16%
14%
12%
10%
8%
6%
4%
2%
0%
Task Orientation
Generating Ideas
Setting Procedural Goals
Setting Content Goals
Setting Rhetorical Goals
Setting Structural Goals
Memorial Pre-text
Reviewing Surface Features
Reviewing Content Goals
Reviewing Rhetorical Goals
Reviewing Structural Goals
Reading Sources
Reading Task
Reading one's own text (part)
Reading one's own text (whole)

Statistics

- Are men and women equally skilled in thinking aloud?
- Is there a difference in the use of mental processes between the two assignments?
- Does the level of ability influence thinking aloud?

Distribution of the total dataset





Exploratory analyses

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	HOD1 of HOD2	Statistic	df	Sig.	Statistic	df	Sig.
Pauzes langer of gelijk aan 2.5 sec	Opdracht 1	.339	819	.000	.332	819	.000
	Opdracht 2	.323	515	.000	.380	515	.000

a. Lilliefors Significance Correction

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Pauzes langer of gelijk aan 2.5 sec	Based on Mean	1.489	1	1332	.223
	Based on Median	.334	1	1332	.564
	Based on Median and with adjusted df	.334	1	1272.946	.564
	Based on trimmed mean	.401	1	1332	.527

Non- Parametric Tests

- Data with non-parametric assumptions (notably normally distributed data)
- Differences between means
- Non-parametric tests→ assumption-free tests
- Principle of ranking the data
- Most common procedures: The Mann-Whitney test, Wilcoxon signed-rank test, Friedman’ s test and Kruskal-Wallis test

Comparing two independent conditions

- Differences between two conditions and different participants in each condition →

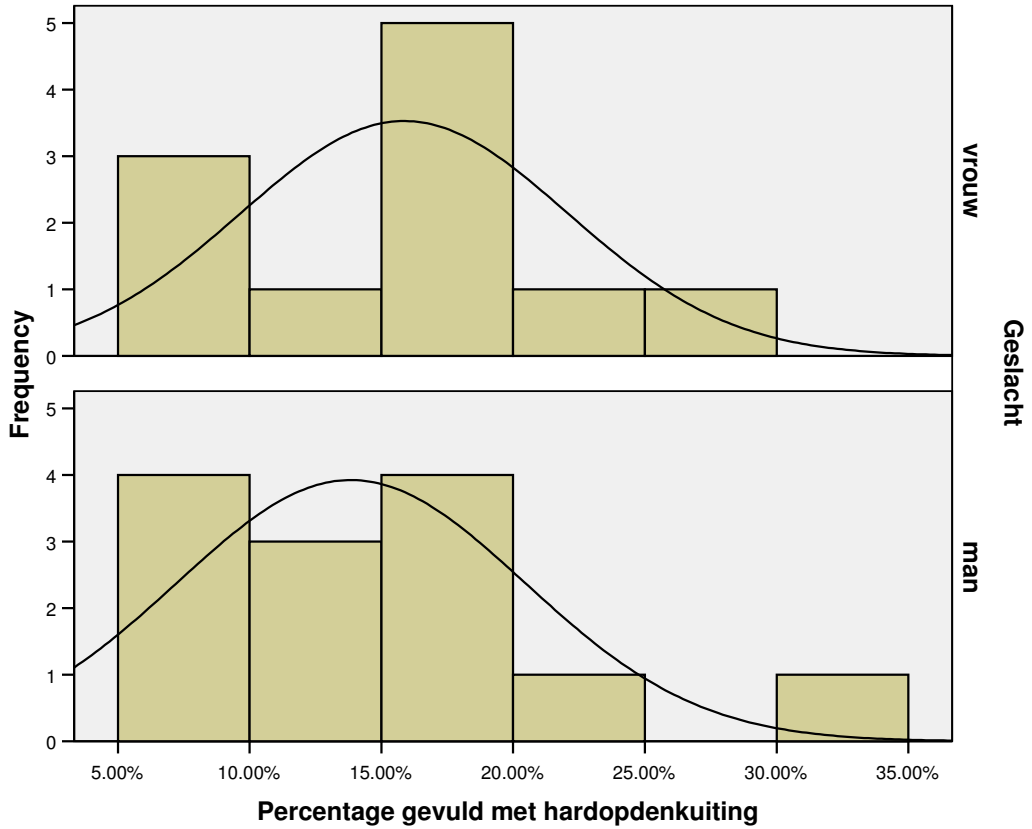
  Mann-Whitney and Wilcoxon rank-sum test

- Equivalent of the independent *t*-test

- Are men and women equally skilled in thinking aloud?

- Is there a difference in the use of mental processes between the two assignments?

Distribution Percentages Men & Women



Ranking the data

Student	Sex	Level of ability	Assignment	Percentage	Rank
15	1	2	1	30.05%	24
3	2	3	1	26.72%	23
11	2	2	1	22.64%	22
24	1	3	2	20.00%	21
6	2	1	2	19.32%	20
2	2	1	2	19.20%	19
18	1	2	1	18.02%	18
21	1	1	2	17.24%	17
9	2	1	1	16.39%	16
5	2	2	2	16.36%	15
23	1	1	1	16.23%	14
4	2	3	2	16.05%	13
14	1	1	1	15.08%	12
1	2	2	1	13.36%	11
16	1	1	1	11.76%	10
19	1	2	2	11.11%	9
20	1	1	2	10.81%	8
8	2	3	2	9.68%	6.5
7	2	3	1	9.68%	6.5
22	1	2	2	7.84%	5
12	1	3	1	7.59%	4
13	1	3	2	7.23%	3
17	1	3	1	6.98%	2
10	2	2	2	5.08%	1

- Sum of Ranks:
- Men:147
- Woman:153

Significance

- Mean (W_s) and standard error (SE_{Ws}) can be calculated from sample sizes of each group:

$$\overline{W_s} = \frac{n_1(n_1 + n_2 + 1)}{2} \qquad \overline{Ws} = \frac{13(13+11+1)}{2} = 162.5$$

$$SE_{Ws} = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}} \qquad SE_{Ws} = \sqrt{\frac{(13 \cdot 11)(13+11+1)}{12}} = 17.26$$

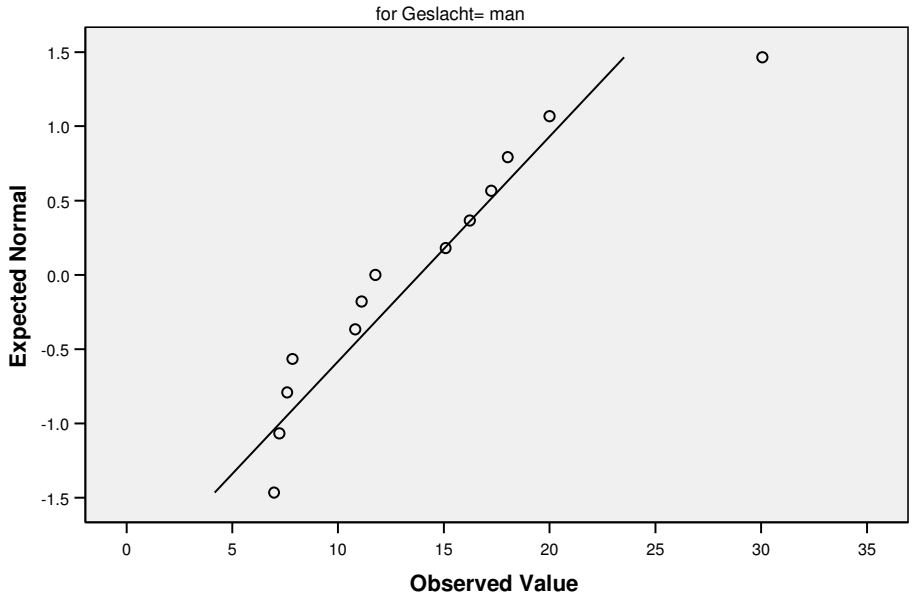
Z-score

$$z = \frac{X - \bar{X}}{s} = \frac{W_s - \bar{W}_s}{SE_{Ws}}$$

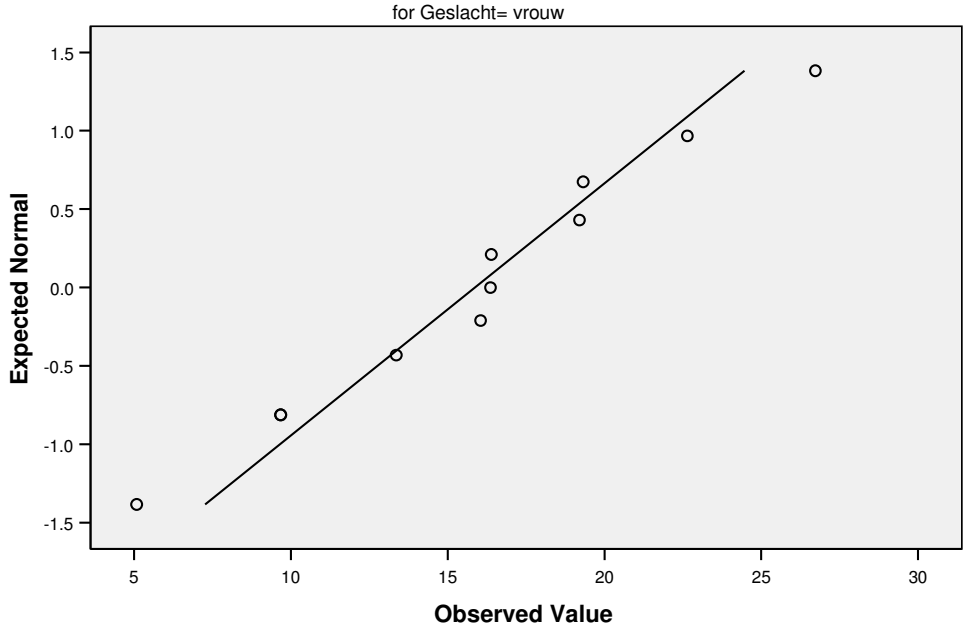
$$z_{man} = \frac{W_s - \bar{W}_s}{SE_{Ws}} = \frac{147 - 162.5}{17.26} = -0.89802$$

$$z_{vrouw} = \frac{W_s - \bar{W}_s}{SE_{Ws}} = \frac{153 - 162.5}{17.26} = -0.5504$$

Normal Q-Q Plot of Percentage gevuld met hardopdenkuiting



Normal Q-Q Plot of Percentage gevuld met hardopden



Exploratory analyses

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Geslacht	Statistic	df	Sig.	Statistic	df	Sig.
Percentage gevuld met hardopdenkuiting	man	.162	13	.200*	.886	13	.086
	vrouw	.148	11	.200*	.979	11	.962

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Percentage gevuld met hardopdenkuiting	Based on Mean	.101	1	22	.753
	Based on Median	.056	1	22	.815
	Based on Median and with adjusted df	.056	1	21.558	.815
	Based on trimmed mean	.083	1	22	.776

Results Sex

- Output Mann-Whitney-test:

Ranks

	Geslacht	N	Mean Rank	Sum of Ranks
Percentage gevuld met hardopdenkuiting	man	13	11.31	147.00
	vrouw	11	13.91	153.00
	Total	24		

Test Statistics[b]

	Percentage gevuld met hardopdenkuiting
Mann-Whitney U	56.000
Wilcoxon W	147.000
Z	-.898
Asymp. Sig. (2-tailed)	.369
Exact Sig. [2*(1-tailed Sig.)]	.392[a]

a. Not corrected for ties.

b. Grouping Variable: Geslacht

Results Sex *t*-test

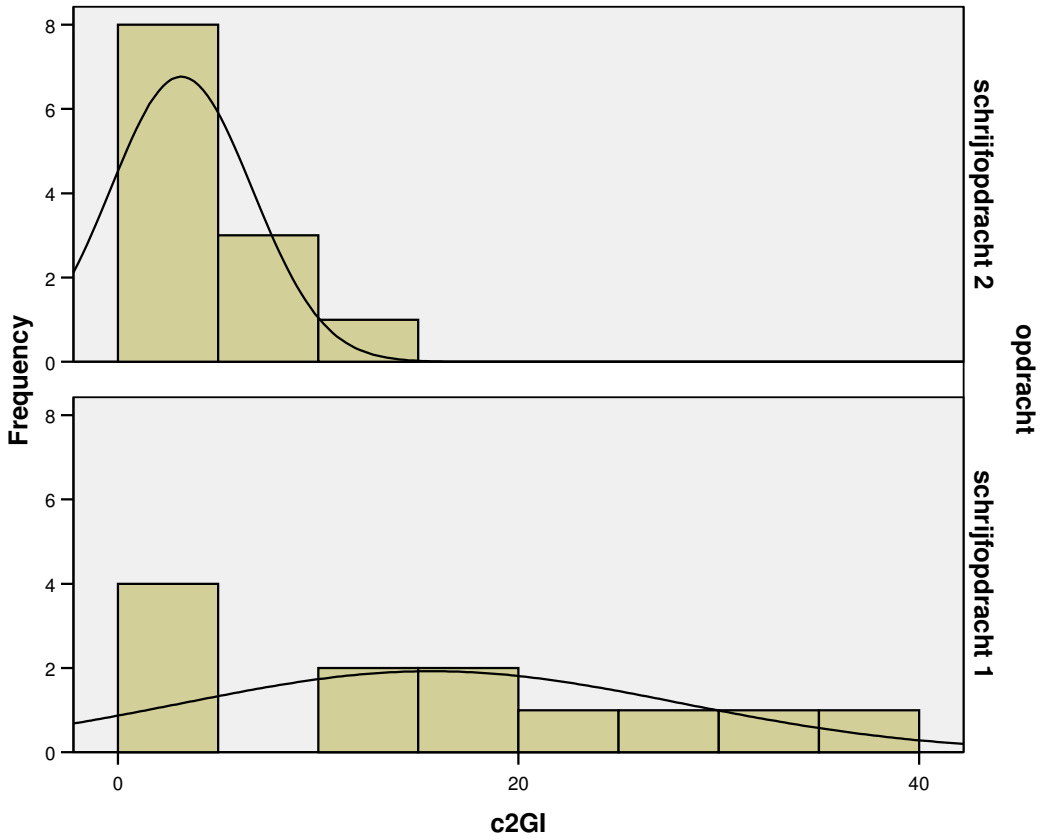
Group Statistics

	Geslacht	N	Mean	Std. Deviation	Std. Error Mean
Percentage gevuld met hardopdenkuiting	man	13	13.8427%	6.60701%	1.83245%
	vrouw	11	15.8629%	6.21692%	1.87447%

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Percentage gevuld met hardopdenkuiting	Equal variances assumed	.101	.753	-.767	22	.451	-2.02024%	2.63528%	-7.4855%	3.44499%
	Equal variances not assumed			-.771	21.718	.449	-2.02024%	2.62136%	-7.4607%	3.42023%

Is there a difference in the use of mental processes between the two assignments?



Exploratory analyses for each group

Tests of Normality

	opdracht	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
c2GI	schrijfopdracht 1	.179	12	.200*	.914	12	.238
	schrijfopdracht 2	.230	12	.080	.833	12	.023

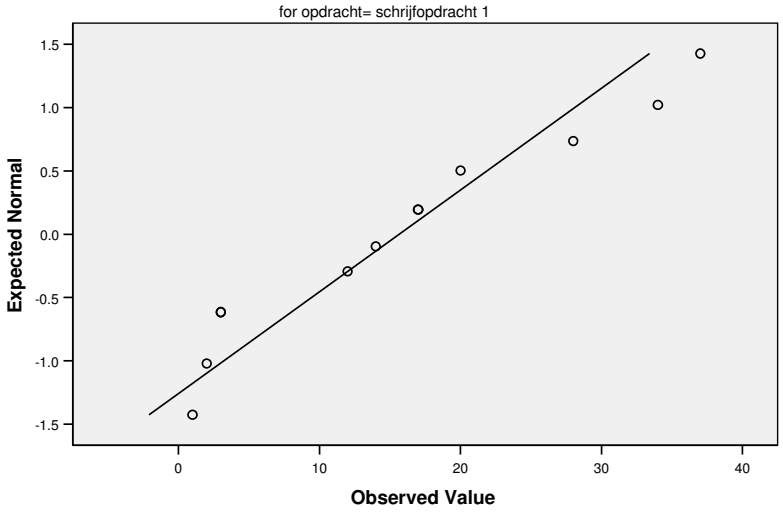
*. This is a lower bound of the true significance.

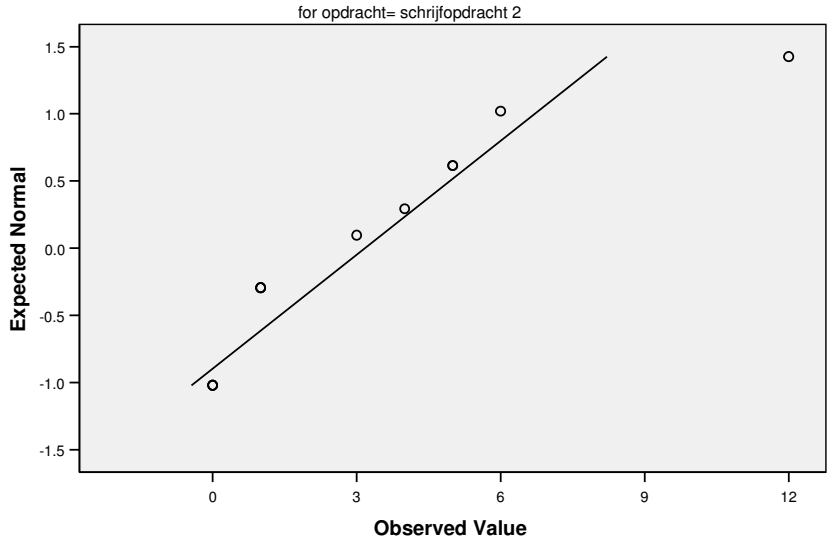
a. Lilliefors Significance Correction

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
c2GI	Based on Mean	11.309	1	22	.003
	Based on Median	11.054	1	22	.003
	Based on Median and with adjusted df	11.054	1	13.730	.005
	Based on trimmed mean	11.309	1	22	.003

Normal Q-Q Plot of c2GI





Which test?

- Two conditions and different participants in each condition
- Difference between the two groups
- Normally distributed data, but heterogeneous variances→ Mann -Whitney

Results Generating Ideas

Ranks

	opdracht	N	Mean Rank	Sum of Ranks
c2GI	schrijfopdracht 1	12	16.42	197.00
	schrijfopdracht 2	12	8.58	103.00
	Total	24		

Test Statistics[b]

	c2GI
Mann-Whitney U	25.000
Wilcoxon W	103.000
Z	-2.726
Asymp. Sig. (2-tailed)	.006
Exact Sig. [2*(1-tailed Sig.)]	.006[a]

a. Not corrected for ties.

b. Grouping Variable: opdracht

Differences between several independent groups: the Kruskal-Wallis test

- Theory Kruskal-Wallis is similar to the Mann-Whitney
- Based on ranked data
- Test statistic, H, is calculated in the following equation:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{K} \frac{R_i^2}{n_i} - 3(N+1)$$

Exploratory analyses for each group

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Niveau	Statistic	df	Sig.	Statistic	df	Sig.
Percentage gevuld met hardopdenkuiting	Hoog	.185	8	.200*	.912	8	.366
	Middel	.131	8	.200*	.972	8	.916
	Laag	.301	8	.031	.830	8	.060

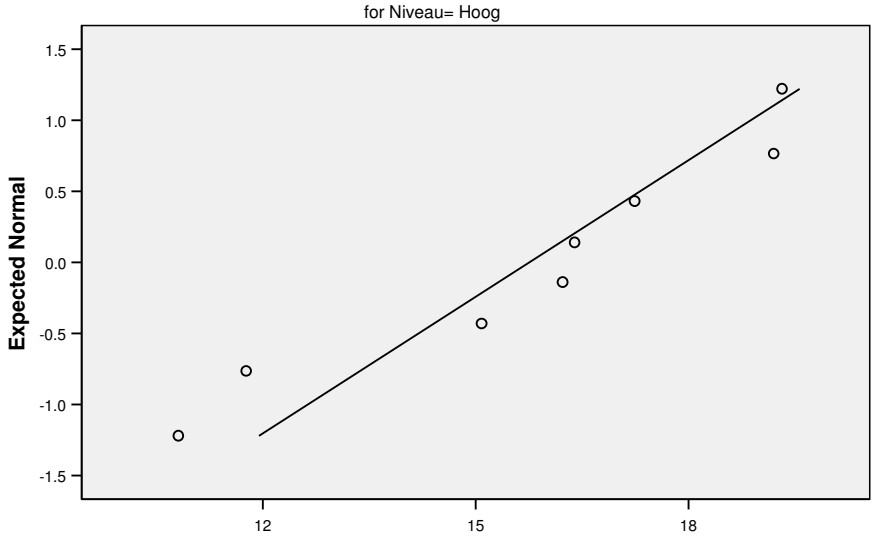
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

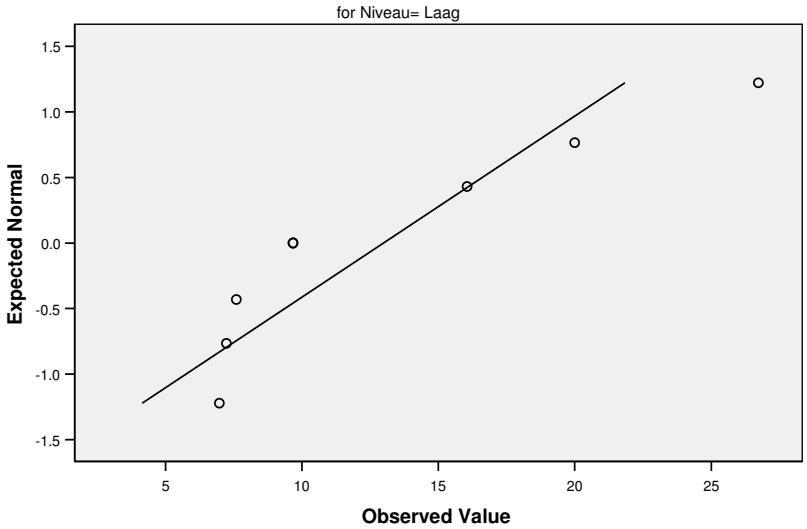
Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Percentage gevuld met hardopdenkuiting	Based on Mean	2.939	2	21	.075
	Based on Median	1.603	2	21	.225
	Based on Median and with adjusted df	1.603	2	15.184	.234
	Based on trimmed mean	2.760	2	21	.086

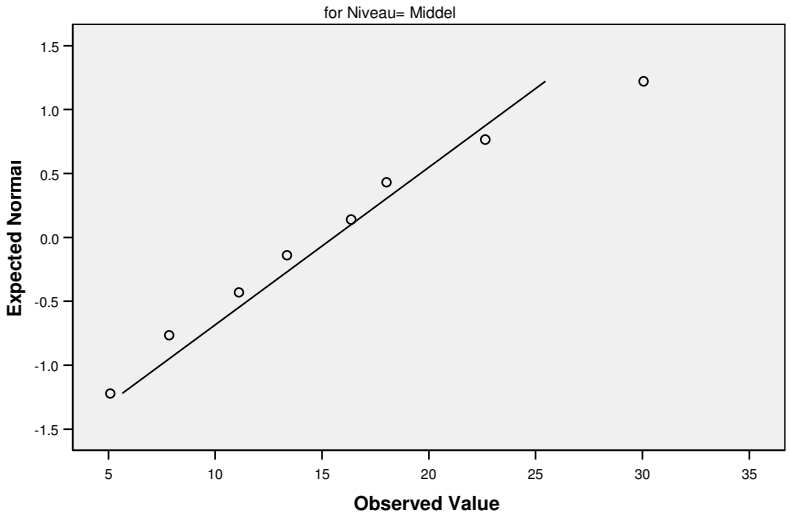
Normal Q-Q Plot of Percentage gevuld met hardopdenkuiting



Normal Q-Q Plot of Percentage gevuld met hardopdenkuiting



Normal Q-Q Plot of Percentage gevuld met hardopdenkuiting



Results level of ability

Ranks

	Niveau	N	Mean Rank
Percentage gevuld met hardopdenkuiting	Hoog	8	14.50
	Middel	8	13.13
	Laag	8	9.88
	Total	24	

Test Statistics[a,b,c]

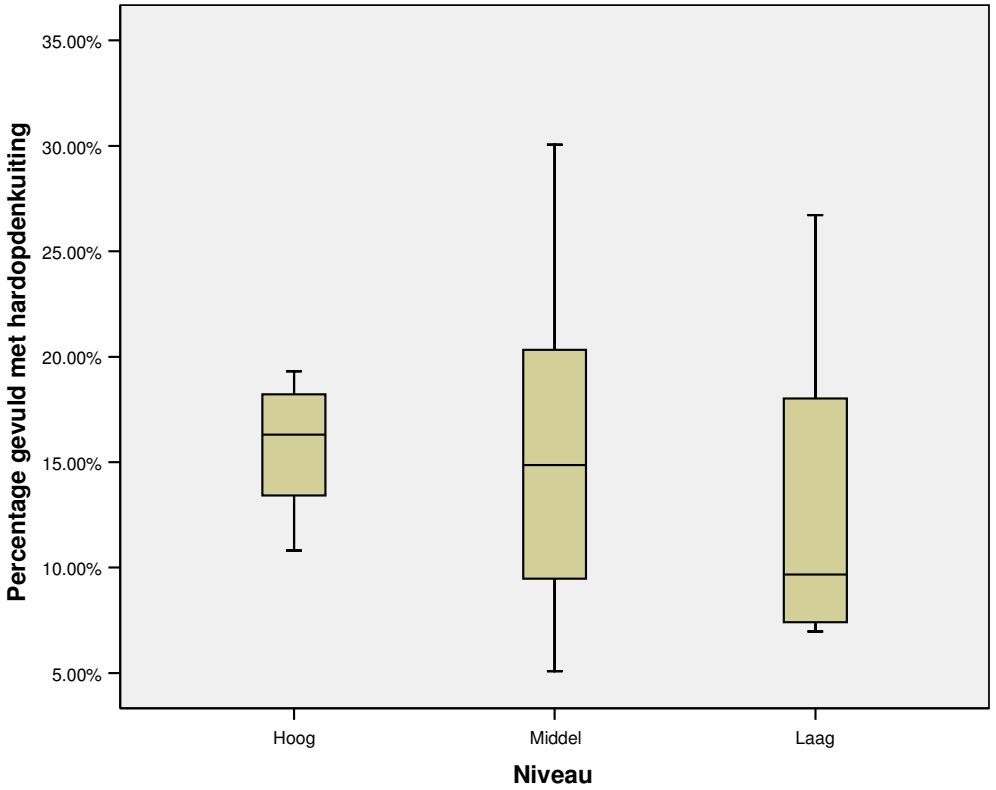
	Percentage gevuld met hardopdenkuiting
Chi-Square	1.806
df	2
Asymp. Sig.	.405

a. Kruskal Wallis Test

b. Grouping Variable: Niveau

c. Some or all exact significances cannot be computed because the time limit has been exceeded.

Do the groups differ?



Post hoc tests

- Mann-Whitney → inflate the Type I error rate
- Mann-Whitney with the *Bonferroni correction*
- Uses critical value of .05 divided by the number of tests you conduct
- .05/3 (level 1→3,1→2,2→3) → level of significance = .0167

Ranks

	Niveau	N	Mean Rank	Sum of Ranks
Percentage gevuld met hardopdenkuiting	Hoog	8	10.13	81.00
	Laag	8	6.88	55.00
	Total	16		

Test Statistics[b]

	Percentage gevuld met hardopdenkuiting
Mann-Whitney U	19.000
Wilcoxon W	55.000
Z	-1.366
Asymp. Sig. (2-tailed)	.172
Exact Sig. [2*(1-tailed Sig.)]	.195[a]

a. Not corrected for ties.

b. Grouping Variable: Niveau

Post hoc test

Ranks

	Niveau	N	Mean Rank	Sum of Ranks
Percentage gevuld met hardopdenkuiting	Hoog	8	8.88	71.00
	Middel	8	8.13	65.00
	Total	16		

Test Statistics[b]

	Percentage gevuld met hardopdenkuiting
Mann-Whitney U	24.000
Wilcoxon W	60.000
Z	-.841
Asymp. Sig. (2-tailed)	.400
Exact Sig. [2*(1-tailed Sig.)]	.442[a]

a. Not corrected for ties.

b. Grouping Variable: Niveau

Test Statistics[b]

	Percentage gevuld met hardopdenkuiting
Mann-Whitney U	29.000
Wilcoxon W	65.000
Z	-.315
Asymp. Sig. (2-tailed)	.753
Exact Sig. [2*(1-tailed Sig.)]	.798[a]

a. Not corrected for ties.

b. Grouping Variable: Niveau

Ranks

	Niveau	N	Mean Rank	Sum of Ranks
Percentage gevuld met hardopdenkuiting	Middel	8	9.50	76.00
	Laag	8	7.50	60.00
	Total	16		

Conclusions

- Thinking aloud not dependent of level of ability or sex.
- Difference between assignment → knowledge telling vs. knowledge transforming

Bibliografie

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Eklundh, K.S, & Kollberg, P. (1996a). Computer tools for tracing the writing process: From keystroke records to S-notation. In G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Eds.). *Theories, models and methodology in writing research* (pp.526-541). Amsterdam: University press.

Eklundh, K.S, & Kollberg, P. (1996b). A computer tool and framework for analyzing online revisions. In C.M. Levy & S. Ransdell (Eds.). *The science of writing. Theories, methods, individual differences, and applications* (pp. 163-188). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Flower, L., & Hayes, J.R. (1980). Identifying the organization of writing processes. In L.W. Gregg & E.R. Steinberg (Eds.). *Cognitive processes in writing* (pp. 3-30). Hillsdale, N.J.: Lawrence Erlbaum Associates Publishers.

Galbraith, D., & Torrance, M. (Eds.). (1999). *Knowing what to write: Conceptual processes in text production.* Amsterdam: University Press.

Gass, S.M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Hillsdale NJ: Lawrence Erlbaum Associated Publishers.

Gregg, L.W., & Steinberg, E.R. (Eds.). (1980). *Cognitive processes in writing*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Keys, C.W. (2000). Investigating the thinking processes of eighth grade writers during the composition of a scientific laboratory report. *Journal of research in science teaching, 37,* 676-690.

Klein, P.D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review*, *11*(3), 203-270.

Levy, C.M., & Ransdell, S. (Eds.). (1996). *The science of writing. Theories, methods, individual differences and applications*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.