



A TAP Study of Thai Test-takers' Processes of Taking a Chinese Vocabulary Test

Xiangdong Gu PhD

Consultant, Cambridge English Language Assessment
Director/Professor, Chongqing University, China

Yuwen Shen

Assistant teacher, Chongqing University, China

11th April, 2014

Contents

- ◆ Think-Aloud Protocols
- ◆ Research Background
- ◆ Research Questions
- ◆ Research Methodology
- ◆ Findings and Discussions
- ◆ Conclusions

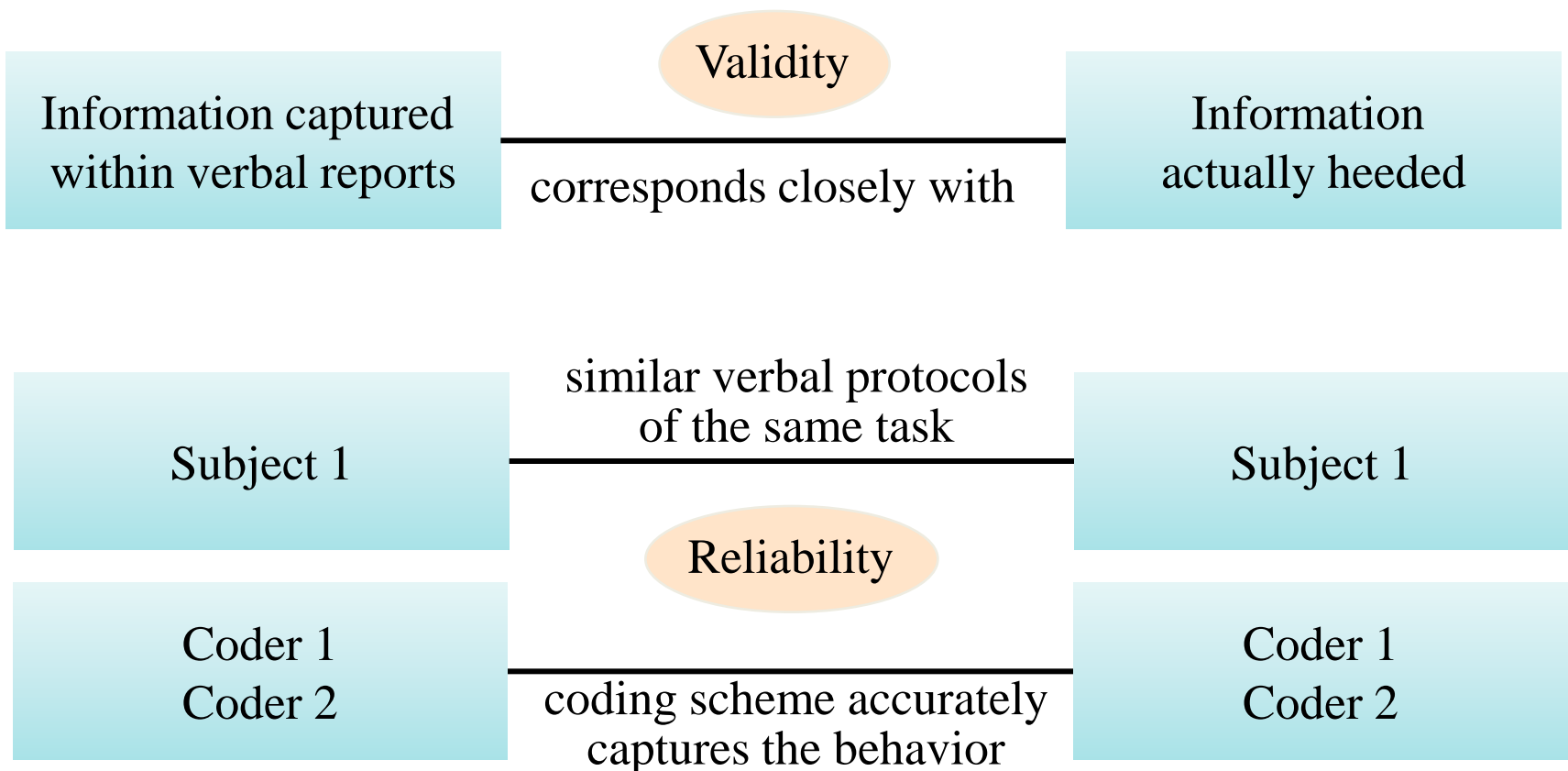
1. Think-Aloud Protocols (TAP)

1.1 Description of TAP

The participants in **think-aloud studies** must perform a particular task while continuously reporting whatever thoughts pass through their working memory. The verbalizations are recorded and then coded according to **theory-driven** and/or **empirically driven** categories.

(Ericsson 1988; Ericsson & Simon 1980, 1993)

1.2 Validity and Reliability of TAP





2. Research Background

2.1 Our Teams' Studies Using TAP

- Gao, X. & Gu, X. (2008) An introspective study on test-taking processes for a **Banked Cloze**, *Teaching English in China* (4), 3-16.
- Zheng, Y. & Gu, X. (2009) A study on the construct validity of **Long Dialogue Multiple-choice Listening Comprehension Tests**: A retrospection approach, *Teaching English in China* (6), 15-26.
- Gu, X. & He, Y. (2011) An introspective study on testees' process of taking **English to Chinese Translation Test** in TEM 8, *Foreign Languages and Translation* (3), 67-73.
- Gu, X. & Shi, C. (2012) A retrospective study on test-takers' cognitive and metacognitive processes in taking a **Compound Dictation Test**, *Chinese Journal of Applied Linguistics* (4), 400-420.
- Gu, X., Lv, Q. & Zhao, L. (2013) A think-aloud protocol analysis of test-takers' processes during the IELTS **Academic Reading Test**, New Directions English: Language Assessment in China, Beijing, 6th Nov. 2013.
- Gu, X., Wang, J., & Xu, J. (2013) A verbal protocol analysis of test-takers' processes of taking CET-4 **Sentence Translation Test**.
- Gu, X., & Shen, Y. (2014) A TAP study of Thai test-takers' processes of taking a **Chinese Vocabulary Test**.



2.2 Brief Introduction to Hanyu Shuiping Kaoshi (HSK)

- A national standardized test launched in **1990** and revised in **2009**
- To assess the **Chinese language proficiency** of **non-native speakers** (including foreigners, overseas Chinese and students from Chinese national minorities)



2.3 The Correspondence Between Levels of HSK and CEFR

HSK (Revised Version)	Vocabulary		CEFR
	Words & Phrases	Characters	
HSK-6	5000+	2663	C2
HSK-5	2500	1685	C1
HSK-4	1200	1064	B2
HSK-3	600	617	B1
HSK-2	300	347	A2
HSK-1	150	174+	A1

(see http://www.hanban.edu.cn/tests/node_7486.html)



2.5 Brief Introduction to HSK-6 Vocabulary Test

- Second part of the reading test
- Altogether 10 short texts
- Each text with 3-5 blanks;
each blank with four options
- Testees required to choose the
most appropriate answer
according to the context

第二部分

第 61-70 题：选词填空。

61. 高中生与初中生相比，在做出判断和决定前能更多地_____各种事实和可能性，_____行动的各种可能后果，决定一旦做出也能更_____地付诸行动。
- | | | | | | |
|------|----|----|------|----|----|
| A 考察 | 预料 | 明显 | B 考虑 | 预计 | 迅速 |
| C 考验 | 预见 | 充分 | D 思考 | 预测 | 显著 |
62. 电影最大的乐趣，在于营造出一个独具_____的光影世界，在这个世界里，不仅可以使人得到_____和娱乐，更可以有所感悟和回味，充分_____到“电影是浓缩的人生”这一特点。
- | | | | | | |
|------|----|----|------|----|----|
| A 魅力 | 放松 | 体会 | B 力量 | 轻松 | 理解 |
| C 特色 | 趣味 | 意识 | D 见解 | 休息 | 反应 |
63. 我们要学会_____自己的心情，而不是让别人决定你的心情，要_____自己对别人坏情绪的“免疫力”，只有这样才能每天_____一份好心情。
- | | | | | | |
|------|----|----|------|----|----|
| A 支配 | 减少 | 得到 | B 控制 | 加强 | 拥有 |
| C 限制 | 增强 | 享受 | D 掌握 | 减弱 | 充满 |
64. 川菜品种_____、味道多变，_____“一菜一格，百菜百味”的美誉。川菜的这种独特的风格也_____国内外人们的青睐，许多人发出“食在中国，味在四川”的_____。
- | | | | |
|------|----|----|----|
| A 丰富 | 享有 | 赢得 | 赞叹 |
| B 多样 | 具有 | 招徕 | 赞赏 |
| C 繁多 | 形成 | 吸引 | 赞扬 |
| D 丰满 | 达到 | 造成 | 赞美 |
65. 喜鹊是自古以来深受人们喜爱的鸟类，是好运与福气的象征。喜鹊登梅是中国画中非常常见的_____，它还经常出现在中国传统诗歌、对联中。此外，在中国的民间_____中，每年的七夕，人间所有的喜鹊会飞上天河，_____起一座鹊桥，让分离的牛郎和织女相会，_____在中华文化中鹊桥常常成为男女情缘的象征。
- | | | | |
|------|----|---|----|
| A 作品 | 节目 | 架 | 因此 |
| B 主题 | 故事 | 洼 | 总之 |
| C 题材 | 传说 | 搭 | 因而 |
| D 材料 | 风俗 | 组 | 于是 |



3. Research Questions

- What types of **information sources** do test-takers use to answer vocabulary test of reading section in HSK-6? How do they use these information sources?
- What types of **strategies** do test-takers employ in their processes of taking vocabulary test? How do they use these strategies?
- Are there any **differences** among different proficient groups and in their use of information sources and strategies? If so, what are these differences?

4. Research Methodology

4.1 Research subjects

➤ Subjects

12 **Thai** students taking the program of Master of Teaching Chinese to Speakers of Other Language in Chongqing University





➤ Three Groups

Group	Gender	Age	HSK Grade & Score (total:300)
H1	M	27	HSK-6 (225)
H2	M	27	HSK-6 (218)
H3	F	28	HSK-6 (214)
H4	F	26	HSK-6 (206)
M1	M	25	HSK-6 (188)
M2	F	26	HSK-6 (192)
M3	F	24	HSK-5 (223)
M4	F	27	HSK-5 (221)
L1	M	25	HSK-5 (207)
L2	F	24	HSK-5 (183)
L3	F	26	HSK-4 (205)
L4	F	25	HSK-4 (211)



4.2 Instrument: Vocabulary test of HSK-6 (April, 2011)

➤ **Directions:** Choose the most appropriate answer according to the context.

61. 相对于年轻人来说，老年人更乐于回忆____。他们不再____获得新的知识，而是专注于昔日的经历，为一生的____做一个总结。

A. 传说 试验 事迹

B. 从前 品尝 时光

C. 现状 摸索 事业

D. 往事 尝试 岁月

61. Compared with young people, the old tend to recall _____. They will no longer _____ acquiring new knowledge but concentrate on their past experiences, and make a summary of their life-long _____.

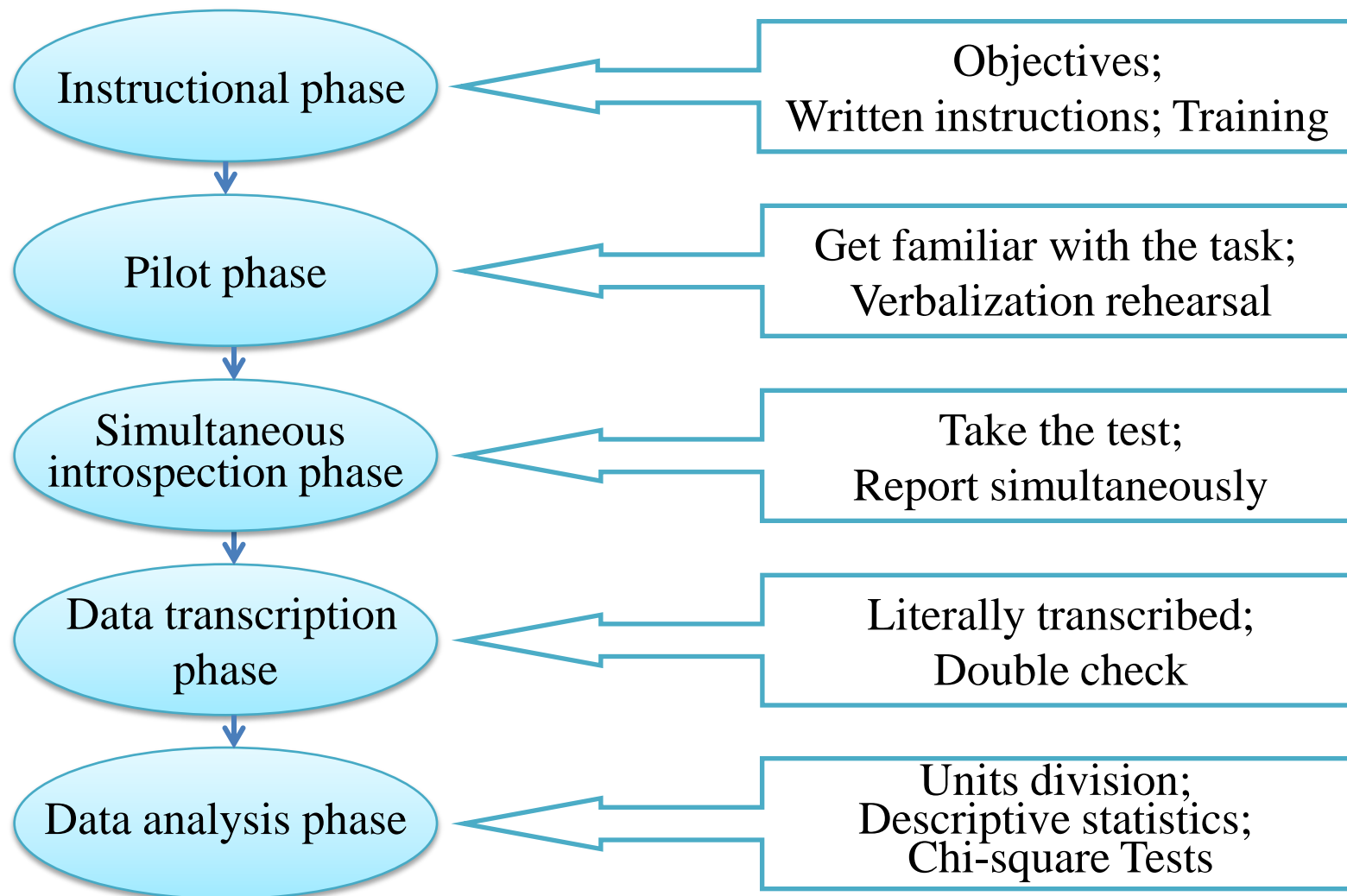
A. **legend**, test, deeds

B. **the past**, taste, time

C. **current situation**, explore, career

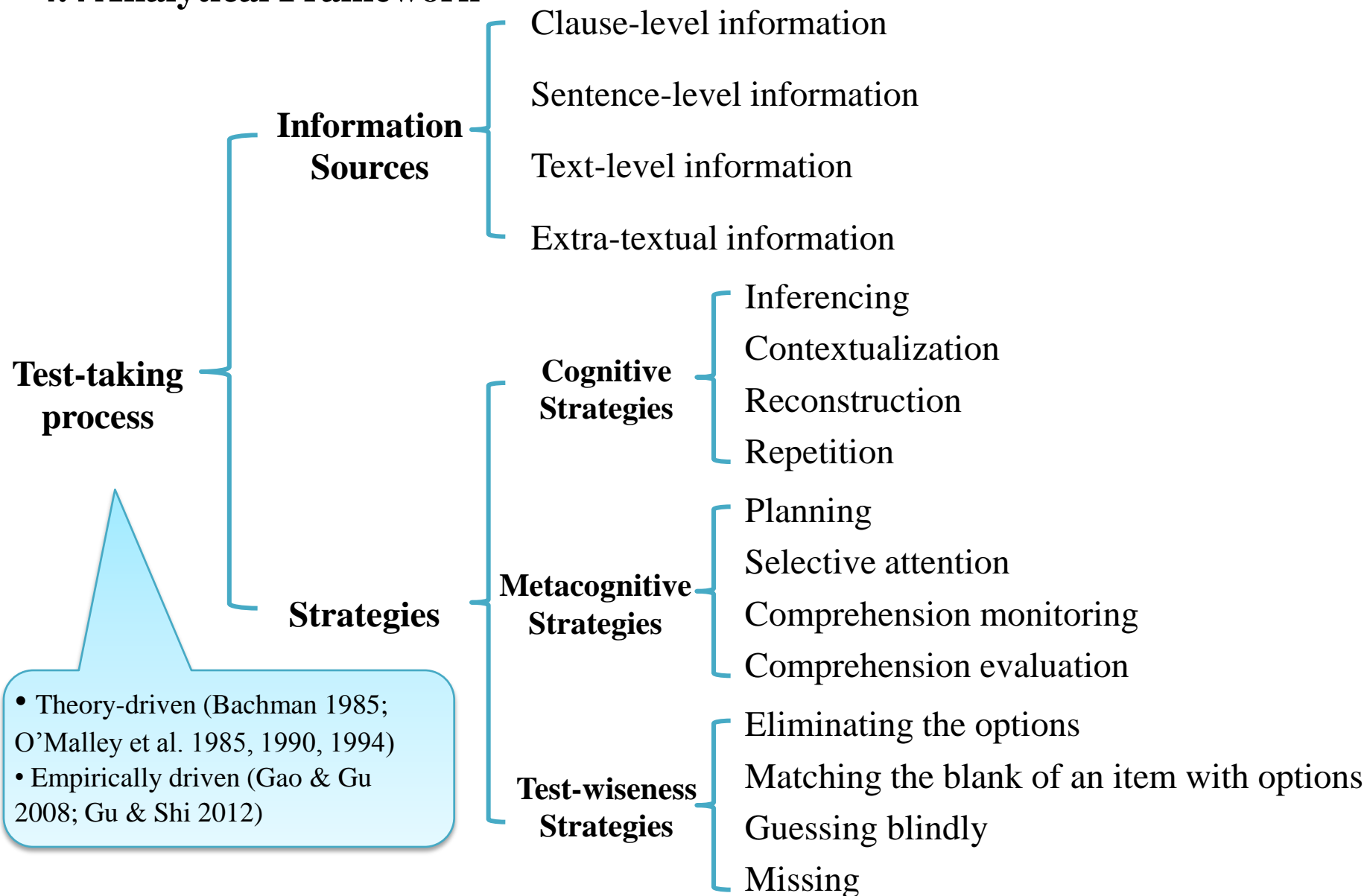
D. **the past experience**, try, years

4.3 Research Procedures



Intra-coder agreement 0.88; Inter-coder agreement 0.81

4.4 Analytical Framework



5. Findings and Discussions

5.1 Information sources

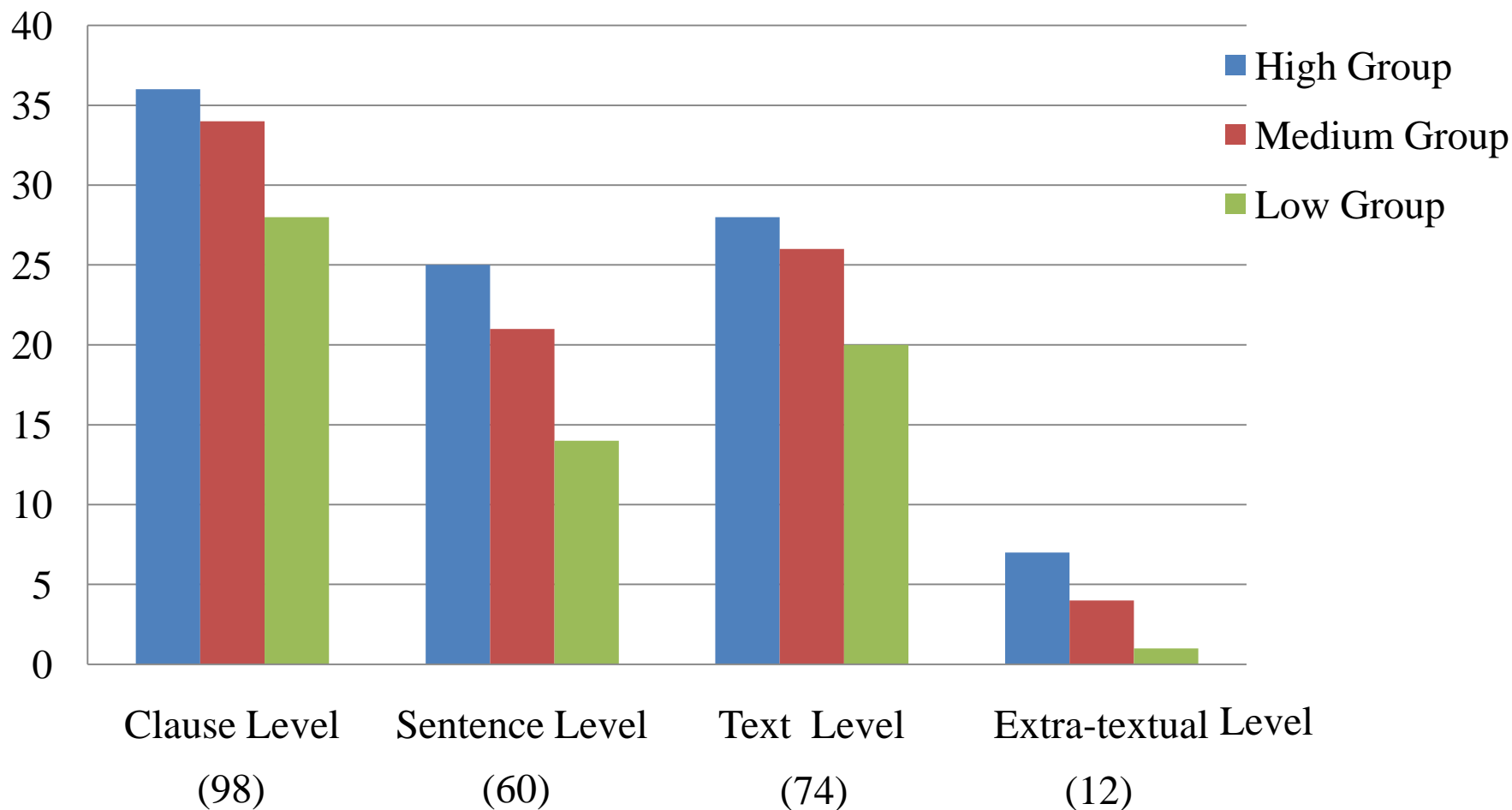


Figure 1 Frequencies of information categories used by the three groups of test-takers



➤ Clause-level information

- The high group used clause-level information in combination with sentence-level and text-level information in most cases.

Example 1 (H1)

因为有乘客时，司机有什么什么，司机有A标志，B目标，C对象，D焦点。这个还不是很清楚，往后看，走到十字路口往左转右转犹豫不定，意识就被分散了，精力就被分散了，活力就被分散了，意志就被分散了。我觉得应该是C，因为有乘客时，司机有对象，他会专心致志于驾驶，小心什么什么的，而没有乘客时，他是冲动的，走到十字路口什么什么，活力就被分散了。

*Because when there are passengers in the car, **the driver has bla bla**. The driver has A. sign B. objective C. target D. focus. **Not clear yet and just go ahead**. Arriving at the crossroad, he hesitates to go left or right, hence awareness/attention/vitality/resolution is distracted. I think C is right. When passengers are in the car, the driver has a target and will focus on driving with caution against bla bla. However, when there are no passengers, he is impetuous. Arriving at the crossroad bla bla, his attention will be distracted.*



- The low group tended to use clause-level information as the single source of information.

Example 2 (L3)

有乘客时，司机有...,目标，对象,司机有对象，不合适，应该是司机有目标.。

*When there are passengers, **the driver has a...target/objective**. The driver has an objective. Not suitable. **It should be that the driver has a target.***



➤ Sentence-level information

- The high group regarded the whole sentence as one unit.

Example 3 (H3)

团购就是团体购物，指的是认识的或者不认识的消费者，联合、结合、联络、合并起来，觉得是.....联合起来，然后再看，来提升买家与商家的议价能力，只有消费者联合起来才能提升买家与商家的议价能力，所以选联合。

Tuangou is short for shopping in group. It refers to consumers known or unknown to combine/join/contact/merge together. I think the answer is “combine together”. Then, to enhance buyers’ bargaining ability with sellers. Only through combining customers together can enhance buyers’ bargaining ability with the sellers. So I choose “combine together”.



- The low group made haste to seek the answers.

Example 4 (L1)

团购，就是团体购物，指的是认识的或者不认识的消费者什么什么起来，起来，联合、结合、联络、合并；认识的或者不认识的消费者什么起来，我觉得是联合起来，所以选A。

Tuangou is short for shopping in group. It refers to consumers known or unknown ...together. combine/join/contact/merge. Consumers known or unknown ...together. I think the answer is “combine together”, so I choose A.



➤ Text-level information

- Every item of the vocabulary test consists of 2-3 sentences with 3-5 words deleted, i.e. a short text. This kind of text generally has centralized information. When test-takers can't solve the problems with the obvious clause-level information, they tend to utilize the text-level information.

Example 5 (H1)

好，第68题，出租车在什么时候最危险？答案是没有乘客时。因为有乘客时，司机有什么什么，司机有A标志，B目标，C对象，D焦点。这个还不是很清楚，他会怎么怎么，他会，迫不及待，全神贯注，这是.....我觉得B不对，与驾驶，什么什么，齐心协力，B想法设法，C小心什么什么，D千方百计；尽快到达目的地，而没有乘客时，他是什么什么的，他是茫然的，他是盲目的，他是抽动的，D他是急躁的，D不对，走到十字路口往左转右转犹豫不定，意识就被分散了，精力就被分散了，活力就被分散了，意志就被分散了。我觉得应该是C，因为有乘客时，司机有对象，他会专心致志于驾驶，小心什么什么的，而没有乘客时，他是冲动的，走到十字路口什么什么，活力就被分散了。



Example 5 (H1)

Well, item 68, When is it the most dangerous for a taxi? for a taxi? The answer is when no passengers are in the car. Because when there are passengers, the driver has bla bla. The driver has A. sign B. target c. objective D. focus. I am not sure yet. He will bla bla. He will, cannot wait for, concentrate on. This...I think B is incorrect, on driving, bla bla, make concerted efforts, B. try every means, C be cautious bla bla, D. make every endeavor to arrive at the destination as soon as possible. When there are no passengers, he is bla bla. He is confused/ blind/impetuous/irritable. D. He is irritable. D is wrong. Arriving at the crossroad, he hesitates to go left or right, awareness/attention/vitality/resolution is distracted. I think C is right. Because when passengers are present, the driver has a target and will focus on driving with caution against bla bla. However, when there are no passengers, he is impetuous. Arriving at the crossroad bla bla, his attention will be distracted.



➤ Extra-level information

- Deriving from the testees' background knowledge

Example 6 (M3)

这样的公司及个人由于太什么，很难有大的发展。根据我的经验和知识，固执就是不太好的吧，就很难有大的发展。所以D为正确答案。

Companies and people like this are too bla to make great progress. According to my experience and knowledge, stubbornness may not be good, and then they hardly make great progress. So D is correct.

5.2 Strategy use

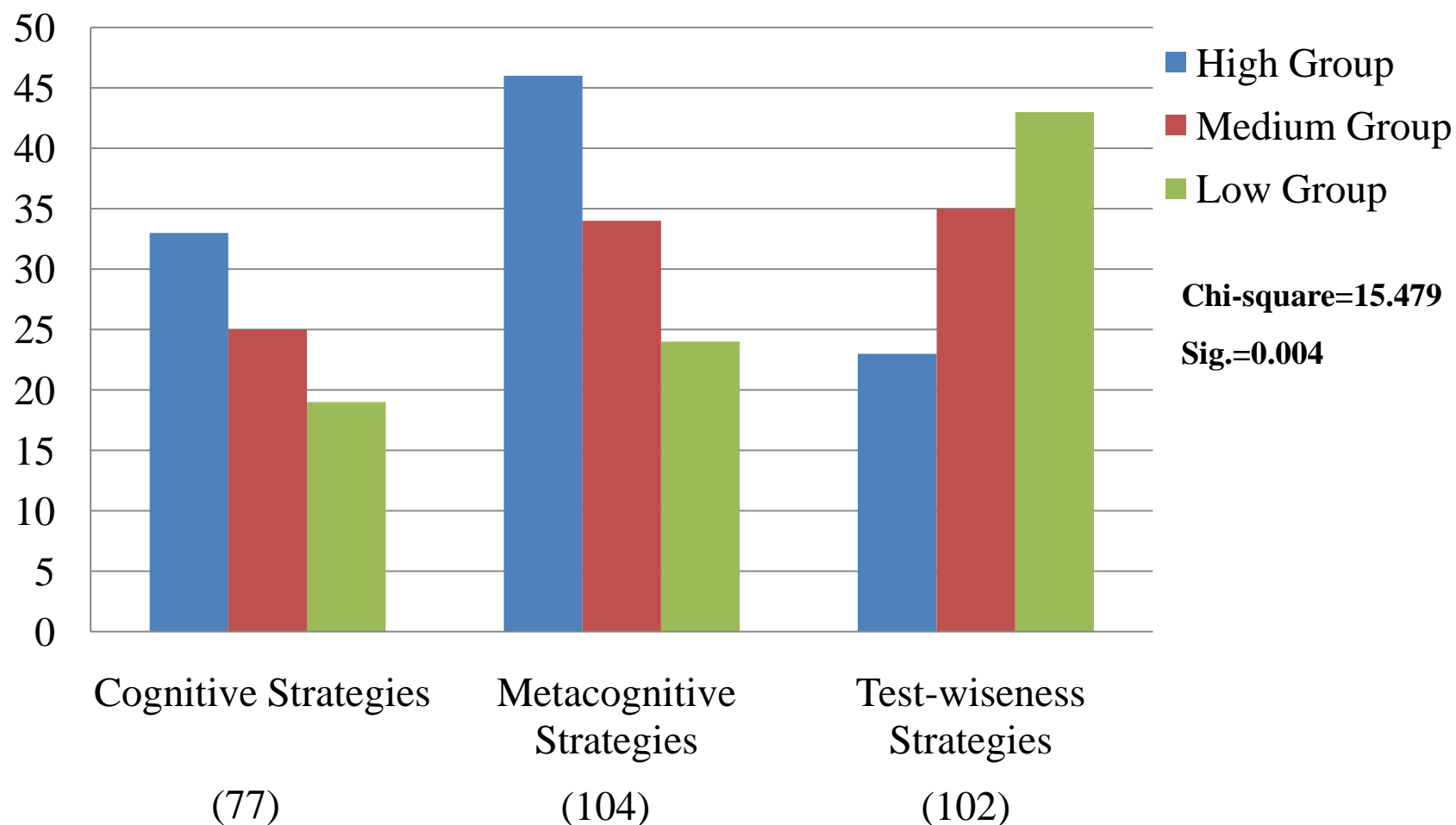


Figure 2 Frequencies of strategy categories used by the three groups of test-takers



Table 1 Cognitive Strategies

Strategy	Description	Protocol
Inferencing	Infer or guess missing information or meaning of words or sentence	相对于年轻人来说，老年人更乐于什么什么，然后去看答案，A传说，B从前，C现状，D往事，这个呢，我觉得先排除A和C，因为老年人更乐于回忆，所以说，更可能说是以前的事情嘛， (Because the old are tend to recall, it is more likely to talk about the past events.) 所以啊，这个呢，应该选B或者选D，从前或者往事。
Contextualization	Relate new information to a wider familiar context	先有目标，他会，看B 的盲目，看不到路，也是看不到，也不知道该怎么做，后面也说了， (Which has been stated in the following part) 走到十字路口往往左转右转犹豫不定，不知道该往哪走，所以他的精力就被分散了。
Reconstruction	Recreate meaning using their own words according to their own understanding	然后我觉得焦点比目标好一些。这个意思就是说 (That is to say) 把乘客当成焦点。然后会聚精会神。然后没有乘客时，精力会被分散。
Repetition	Repeat some words or sentences	是因为消费者 联络，联络(combine, combine) 起来，因为认识的或者不认识的得需要联络起来，以求得最优的一种购物手段，手段 (means, means)?

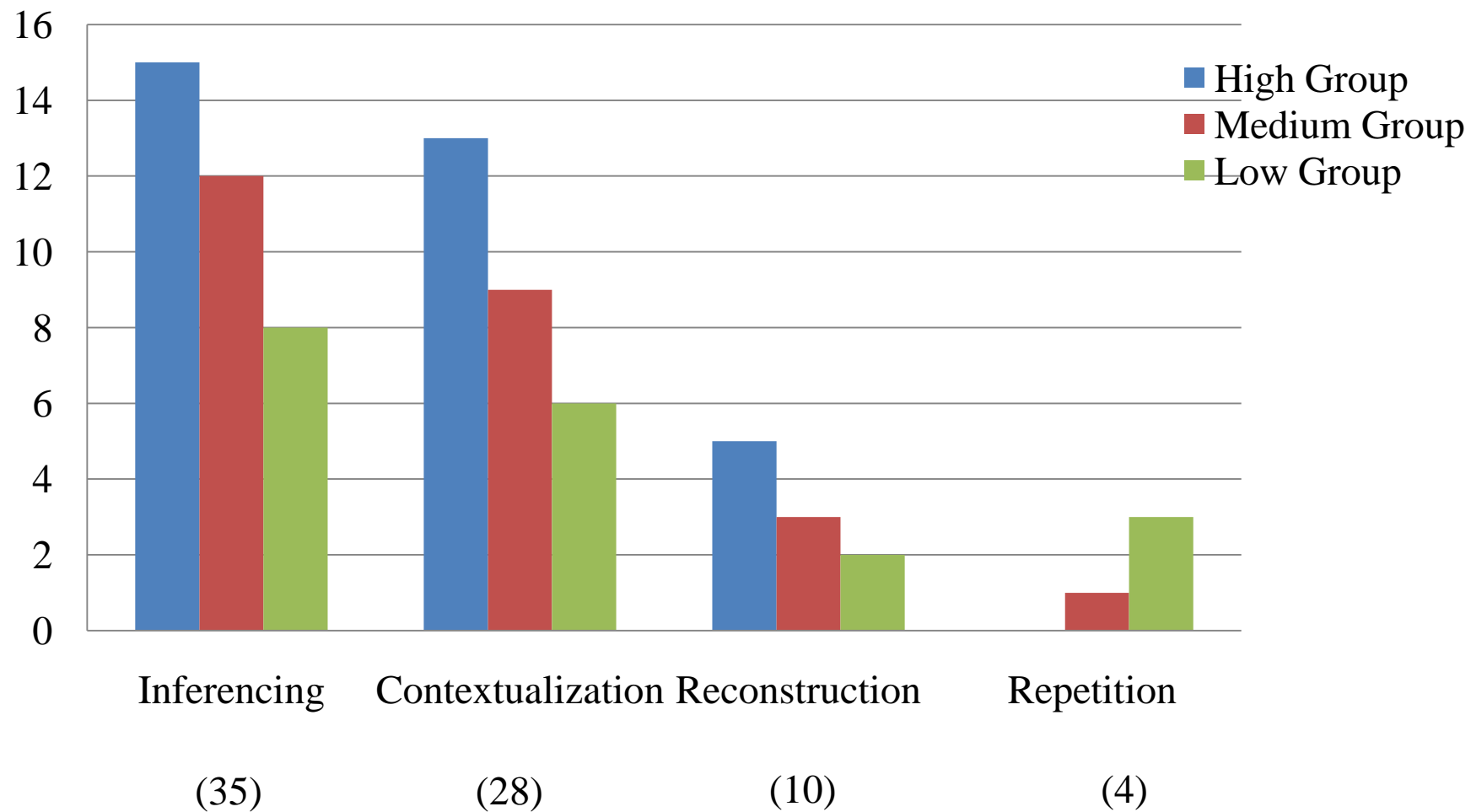


Figure 3 Frequencies of **cognitive strategies** used by the three groups of test-takers



Table 2 Metacognitive Strategies

Strategy	Description	Protocol
Planning	Preview contents in different forms; establish purpose for reading	相对于年轻人来说，老年人更乐于回忆什么，他们不再什么获得新的知识，而是专注于昔日的经历，为一生的什么做一个总结（一个读题的过程，明确目标）(A process of reading the text, in order to understand the requirments)。
Selective attention	Notice specific aspects of reading materials; reading selectively according to purpose	第63题，我答B，是因为，这些选项里我只知道万一的意思(Of all the options I only know the meaning of in case)，把万一放进去，万一制定了基本方针，所以我就答B。
Comprehension monitoring	Monitor understanding while reading	老年人更乐于回忆，该是往事吧，这个意思我不太懂它的意思，先不做，往后猜 (without choosing this blank and going on guessing the following ones)。
Comprehension evaluation	Check understanding for accuracy, completeness and acceptability	第63题，我选择也是D，因为我觉得，一旦制定了基本方针，就一步都不打算妥协，我觉得这样比较顺，别的选择不太顺，所以我选择D。再看其他的空 (and then consider other blanks)，这样的公司及个人由于太固执，很难有大的发展。固执就是不太好的吧，就很难有大的发展。所以我确定D为正确答案 (So I choose D as the right answer)。

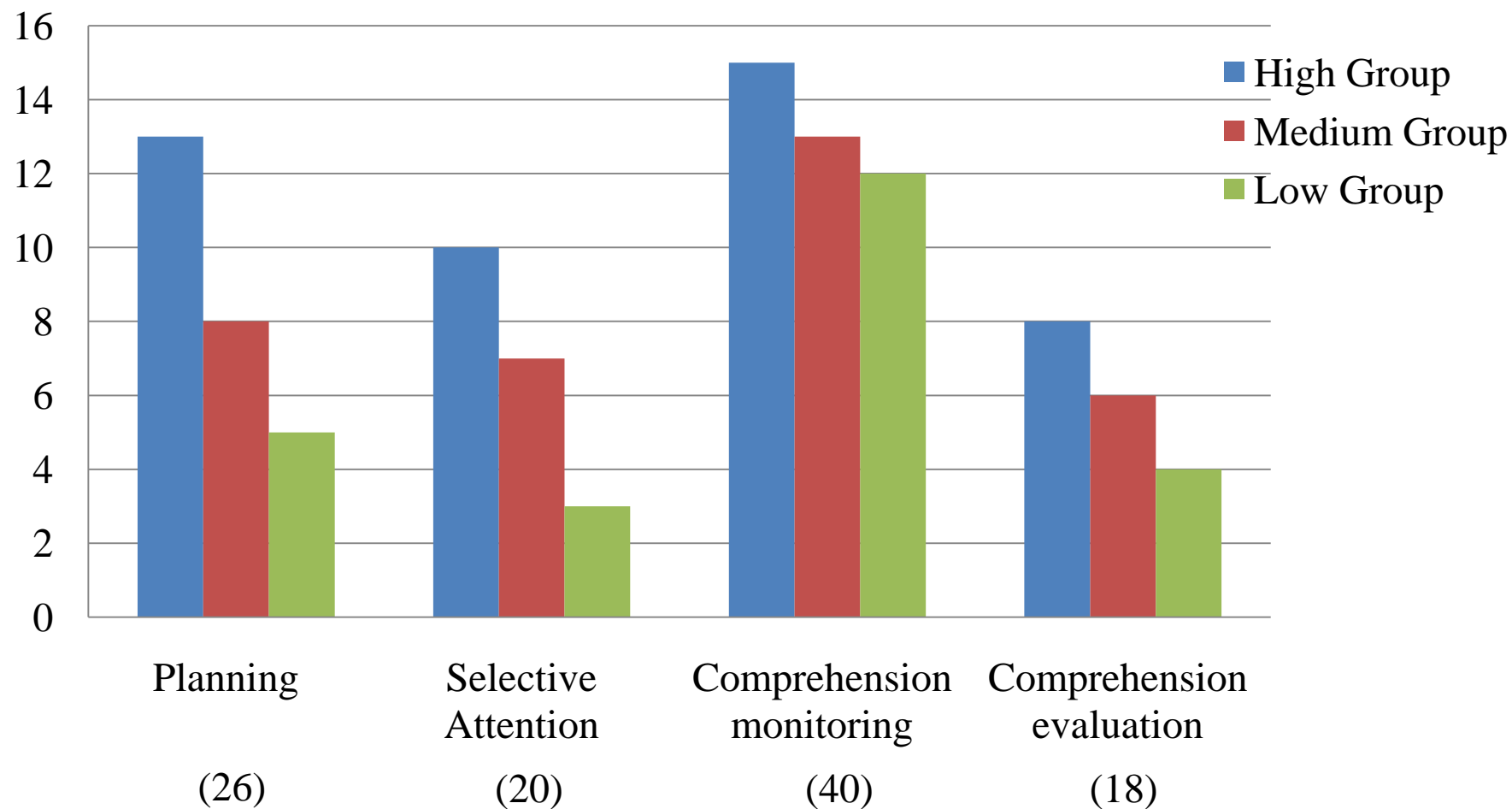


Figure 4 Frequencies of **metacognitive strategies** used by the three groups of test-takers



Table 3 Test-wiseness Strategies

Strategy	Description	Protocol
Missing	Do nothing about some blanks, relying the choices of other blanks	ABC最后的词我不知道什么意思，然后我再看，做生意时总会有一些企业或个人只知道坚持自己的立场，什么什么制定了基本方针，就什么什么，后面有就，前面应该，可能万一，一旦，一旦.....就，应该是一旦.....就，后面就不管了(The following blanks will not be considered.), 选D。
Eliminating the options	Select an answer not because it is thought to be correct, but because the others do not seem reasonable or are not understandable	相对于年轻人来说，老年人更乐于什么什么，然后去看答案，A传说，B从前，C现状，D往事，这个呢，我觉得先排除A和C (In my opinion, A and C could be firstly get rid of.), 因为老年人更乐于回忆，
Matching the blank of an item with the options	Put the given options one by one in the blank in order to see whether it is suitable in the context	出租车在什么时候最危险？答案是没有乘客时。因为有乘客时，司机有什么什么，司机有 A标志，B目标，C对象，D焦点(A. sign, B. destination, C. object, D. focus) 。这个还不是很清楚，他会怎么怎么，他会，迫不及待，全神贯注，我觉得B不对，与驾驶，什么什么，齐心协力,B 想方设法，C小心什么什么，D 千方百计 (B. try every means to, C. be cautious about bla bla, D. make every effort to)。
Guessing blindly	Randomly choose from the given options without any particular consideration	这个63题，答案的词我不太明白，不知道它的意思，然后看起来，我觉得，额，它的答案就是C，这道题我就是猜的 (I worked out this question by mere guessing)

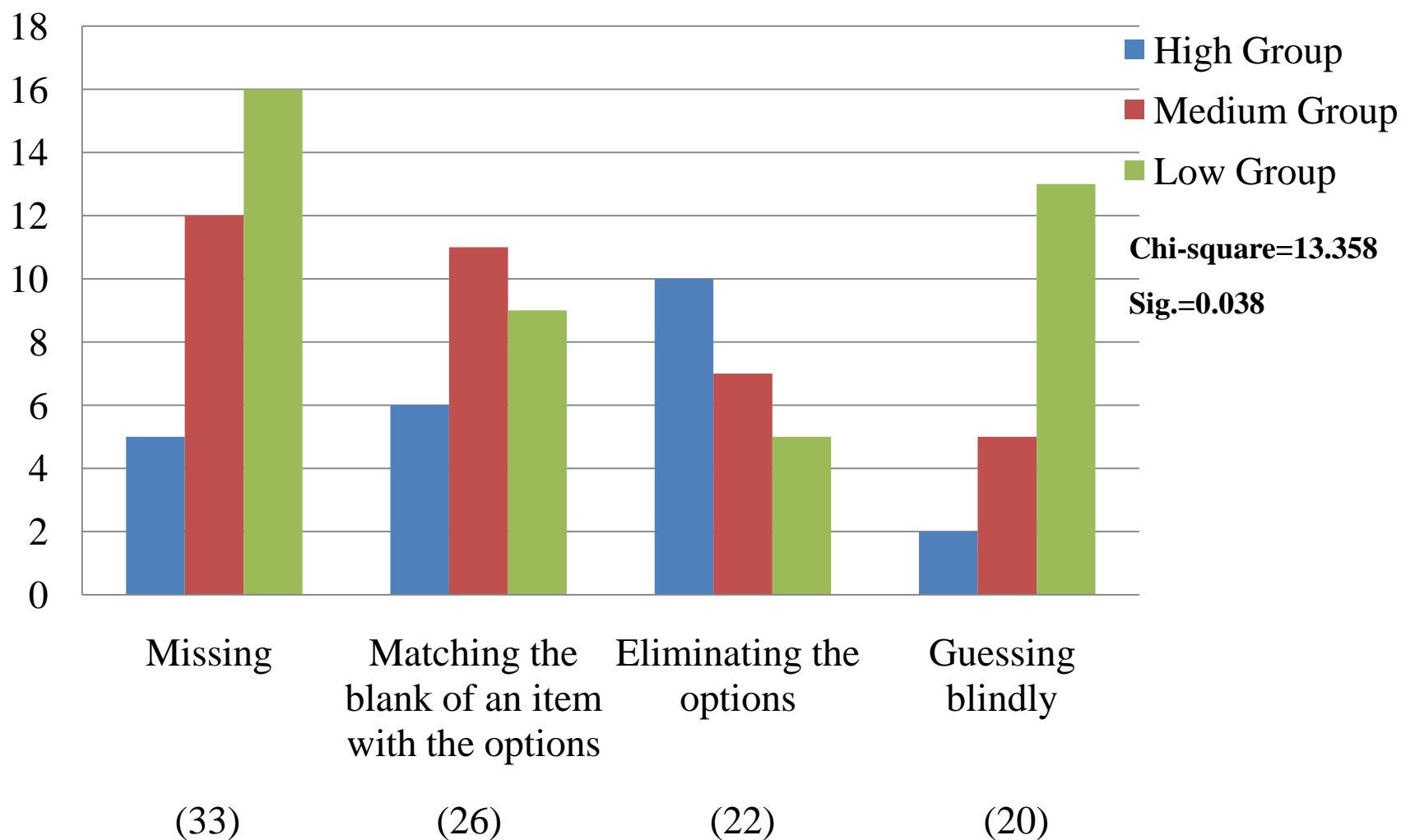


Figure 5 Frequencies of **test-wiseness strategies** used by the three groups of test-takers



6. Conclusions

6.1 Major Findings

Q1: *What types of **information sources** do test-takers use to answer vocabulary test of reading section in HSK-6? How do they use these information sources?*

- Test-takers employ all four categories of information sources, i.e. clause-level, sentence-level, text-level and extra-textual information, among which **clause-level information** is the most frequently used one, followed by text-level, sentence level and extra-textual information.



Q2: *What types of **strategies** do test-takers employ in their processes of taking vocabulary test? How do they use these strategies?*

- Test-takers employ **all three** strategy categories, i.e. cognitive, metacognitive and test-wiseness strategies.
- Among the four cognitive strategies, **inferencing and contextualization** are much more frequently used than reconstruction and repetition.
- In terms of the four metacognitive strategies, **comprehension monitoring** is most frequently used, followed by planning, selective attention and comprehension evaluation.
- As far as the four test-wiseness strategies are concerned, **missing** is used more frequently than matching the blank of an item with the options, eliminating the options, and guessing blindly.



Q3: *Are there any **differences** among different proficiency groups in their use of information sources and strategies? If so, what are these differences?*

- There exists some differences among different proficiency groups in their use of information sources and strategies, but **significant difference** only exists in the **use of test-wiseness strategies**.
- The higher the proficiency of the group, the more information sources it uses, the more information from different levels it combines and the wider range of discourse it incorporates.
- The higher the proficiency of the group, the more cognitive and meta-cognitive strategies it uses. On the contrary, the lower the proficiency of the group, the more test-wiseness strategies it uses.

6.2 Limitations

- The small item number of the instrument
- The small size of the participants
- The unrepresentativeness of the participants from a homogeneous cultural and linguistic background
- Thus the generalizability of the research findings remains to be further testified by triangulation methodology



6.3 Implications

- The text should be lengthened to provide more information and context for the test-takers.
- Care should be taken to avoid activating test-takers' extra-textual information and to avoid their use of test-wise strategies to get the correct answers.
- More training and practice is needed for test-takers to improve their appropriate use of information sources and strategies.



6.4 Significance

- **Theoretically**, offered some insights into what the Chinese vocabulary test measures.
- **Practically**, put forward suggestions for the test designers on test development and test evaluation, and provided implications for language teachers and learners on how to improve Chinese language proficiency.
- **Methodologically**, proved think-aloud protocols as an effective approach to examine test-takers' use of information sources and strategies.



References

1. Bachman, L. F. and A. S. Palmer. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
2. Ericsson, K. A. (1988). Concurrent verbal reports on text comprehension: a review. *Text*, 8 (4), 295-326.
3. Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87 (3), 215-251.
4. Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data* (revised). Cambridge: MIT Press.
5. Gao, X., & Gu, X. (2008). An Introspective Study on Test-taking Process of for Banked Cloze. *CELEA Journal (Bimonthly)*, 31 (4), 3-16.
6. Gu, X., & Shi, C. (2012) A retrospective study on test-takers' cognitive and metacognitive processes in taking a Compound Dictation Test, *Chinese Journal of Applied Linguistics* 35 (4), 400-420.
7. O'Malley, J. M., & Chamot, A.U. (1990). *Learning strategies in second language acquisition*. NY: Cambridge University Press.
8. 国家汉办/孔子学院总部 (2010). 《新汉语水平考试大纲》. 北京: 商务印书馆.

Acknowledgements

- Organizers
- My MA team





Thank you

Contact
information



0086-13983732737



0086-23-65102602



xiangdonggu@263.net