

【原著】

一般化可能性理論による面接データ解析を行うことが可能な面接官の配置について

木村拓也（長崎大学）

本研究の目的は、一般化可能性理論を用いた AO 入試における信頼性評価の方法を提案した木村・吉村(2010)を受け、実際の面接現場に即して、一般可能性理論による面接データ解析を行うことが可能な面接官配置についての事例を提示することである。本研究で提示した事例を活用して面接官配置を決定することにより、各入試で課される主観的評価の 1 つである面接試験の採点データを事後に検証し、次年度に向けて改善を行った上で入試を実施する PDCA サイクルが導入可能となり、面接試験を含んだ入試の信頼性向上が期待できる。

1 問題の所在—大学入試場面における一般化可能性理論の適用可能性

大学入試が多様化して以降、面接試験や小論文試験が各大学の入学者選抜においてより一層重要な位置を占めるようになって久しい。勿論、一般入試の前期・後期日程においても、面接試験や小論文試験などの主観的評価を用いる選抜場面は数多い。だが、その信頼性評価におけるテスト理論の適用は、まだ途についたばかりであり、今後、一層の事例蓄積が求められるであろう。木村・吉村(2010)で紹介したように、本学では AO 入試の面接・課題作文の採点データに一般化可能性理論 (Brennan 2001, 池田 1994, Shaverson & Webb 1991) を適用し、その結果を学部で FD という形で還元することによって、入試改善を行ってきた。

一般化可能性理論には、分散分析の手法を利用して、各要因の大きさ(分散成分)を推定する Generalizability Study(G 研究)と、G 研究で得られた分散成分から各テストデザインの信頼性を評価する Decision Study (D 研究)とがある(池田 2007)。分かりやすく言えば、前者については、面接評価や小論文

採点において、いかなる誤差要因が大きく占めているのかが分かり、後者については、何人の採点者(面接官)で、いくつの観点で試験を行うと、テストとしての信頼性が担保されるのかについての解を得られる(木村・吉村 2010:82)。

実際に、FD を行った成果としては、各学部の入試担当者間で、G 研究の結果から、主観的評価における誤差の認識が深まったことや、どういった観点や質問項目で識別力が高いのかという経験の蓄積・情報共有が可能になったことが挙げられる。何より、エビデンスに基づいて、次年度に課す観点や質問項目の取捨選択についての建設的な議論が深まったことが最大の収穫であった。また、D 研究の結果から、具体的な信頼性指標をもとに多くの人数が関わる AO 入試や推薦入試において、採点者数の削減指針が得られたことも大きい成果であった。

だが、一般化可能性理論の適用に際しては、各学部の全ての面接・課題作文の採点データで可能であった訳ではない。例えば、多くの教官の目で見ることが大切との認識のもとで、面接の途中で、面接官の半数を面接室

間で交互に入れ替えるなどの措置を行った学部については、一般化可能性理論が適用可能な実験計画を見つけることができなかった。

そこで、実際に、各入試場面で課された主観的評価の採点結果を事後に検証し、次年度に向けて改善を施して、入試を実施するPDCAサイクルを導入して、入試の信頼性を向上するために、一般化可能性理論が適用可能な面接官の配置について、一度、事例集という形で周知する必要があると考えた。そこで、本研究の目的を実際の入試現場、特に、面接の場面に即して、客観的にテストとしての評価を行うことが可能な面接官の配置についての事例を提示することに設定したい。

2 面接官配置の問題

具体的に、主観的評価の1つである面接試験について考えてみたい。例えば、表1は、

仮に、朝の10時から始めて夜の6時ちかくまで、面接と面接の間の受験生の移動時間を5分、昼休憩を1時間、約2時間ごとの面接ごとに10分間の休憩をとって面接した場合の仮想スケジュールである。それぞれ、面接時間を15分、20分、30分、40分とった場合に1日で何人面接ができるのかが分かる。それによると、面接時間15分の場合21人、20分の場合19人、30分の場合12人、40分の場合9人である。15分や20分という時間は、挨拶をして志望理由を一通り聞いてしまえば、1つ2つの質問はできても、あつというまに終わってしまう時間である。そうした限られた時間であるにもかかわらず、例えば、AO入試における面接で高校時代に行ってきた研究発表も含めて面接試験を実施するということになれば、更に面接できる人数は限られてしまう。このことを

表1 仮想の面接スケジュール

面接時間 人数	15分	人数	20分	人数	30分	人数	40分
1	10:00 ~ 10:15	1	10:00 ~ 10:20	1	10:00 ~ 10:30	1	10:00 ~ 10:40
2	10:20 ~ 10:35	2	10:25 ~ 10:45	2	10:35 ~ 11:05	2	10:45 ~ 11:25
3	10:40 ~ 10:55	3	10:50 ~ 11:10	3	11:10 ~ 11:40	3	11:30 ~ 12:10
4	11:00 ~ 11:15	4	11:15 ~ 11:35	4	11:45 ~ 12:15		休憩(1時間)
5	11:20 ~ 11:35	5	11:40 ~ 12:00		休憩(1時間)	4	13:10 ~ 13:50
6	11:40 ~ 11:55		休憩(1時間)	5	13:15 ~ 13:45	5	13:55 ~ 14:35
	休憩(1時間)	6	14:00 ~ 14:20	6	13:50 ~ 14:20	6	14:40 ~ 15:20
7	12:55 ~ 13:10	7	14:25 ~ 14:45	7	14:25 ~ 14:55		休憩(10分)
8	13:15 ~ 13:30	8	14:50 ~ 13:15	8	15:00 ~ 15:30	7	15:30 ~ 16:10
9	13:35 ~ 13:50	9	13:20 ~ 13:40		休憩(10分)	8	16:15 ~ 16:55
10	13:55 ~ 14:10	10	13:45 ~ 14:05	9	15:40 ~ 16:10	9	17:00 ~ 17:40
11	14:15 ~ 14:40		休憩(10分)	10	16:15 ~ 16:45		
12	14:45 ~ 15:00	11	14:15 ~ 14:35	11	16:50 ~ 17:20		
	休憩(10分)	12	14:40 ~ 15:00	12	17:25 ~ 17:55		
13	15:10 ~ 15:25	13	15:05 ~ 15:25				
14	15:30 ~ 15:45	14	15:30 ~ 15:50				
15	15:50 ~ 16:05	15	15:55 ~ 16:15				
16	16:10 ~ 16:25		休憩(10分)				
17	16:30 ~ 16:45	16	16:25 ~ 16:45				
18	16:50 ~ 17:05	17	16:50 ~ 17:10				
	休憩(10分)	18	17:15 ~ 17:35				
19	17:15 ~ 17:30	19	17:40 ~ 18:00				
20	17:35 ~ 17:50						
21	17:55 ~ 18:10						

考えてみても、面接試験では、それほど人数がさばけないことが分かる。

となると必然的に、複数の面接室が設置される。だが、違う受験生を前に、同時に複数の面接が行われる状況では、当然のことながら、入試現場では綿密な評価訓練が行える時間的な余裕はないので、例え、打ち合わせがあったとしても、部屋ごとの評価基準が少なからずずれてしまうことは否めない。そこで、そうした誤差がどれほど得点に影響をあたえているのかを評価したり、より誤差が生じにくい観点・質問項目を導入したりして、入試の信頼性を高める努力が求められてくる。こうした信頼性評価は、主観的評価を重視する入試においては喫緊の課題となろう。

勿論、課題論文（小論文）の採点でも、大量の受験者の採点を行う場合、複数の採点者が別々の受験者の採点を行うこともあり、面接官配置と同じ複雑な採点状況が生まれてくる。本稿では、無用の混乱を避けるため、主観的評価の中でも、敢えて、面接試験場面を想定した用語を用いて論じることとする。

3 面接官配置の事例

図1～14は、Brenann(2001)のTable 9.2で示された実験計画を、具体的な面接場面に即して書き直したものである。vはvisionの略で「観点」、iはitemの略で「質問項目」を、pはpupileの略で「受験生」を、hはhumanの略で「面接官」、cはclassroomの略で「面接室」を表している。また、デザインに関する記号の表記については、○および：はネストされた状況を、●および×は、クロスした状況を表すBrenann (2001)のものに従っている(Brenann 2001: 274f)。図示した場合、楕円が包含されているときにはネストされた状況を、楕円が交差しているときにはクロスした状況を表している。

楕円については3要素以上のモデルで表しているが、もちろん、2要素のものでもデータ分析をおこなうことは可能である。楕円が点線であるときには、その要素がなくても分析には問題がない場合である。3要素の図1～4のときには1つが点線で表されており、4要素の図5～14のときには1つが点線で表されており、点線を略した場合、図1か図3のパターンに収斂する。但し、略された要素や他の要素と略された要素との交互作用に関する誤差成分は分析では明らかにはならない。

図1は、全ての受験者に対して質問ごとに異なる面接官集団が採点を行う場合である。例えば、アドミッション専任教員と学部教員が共同で面接したり、学部教員内で役割分担をしたりで、評価する質問項目が異なる場合には、この面接官配置が当てはまる。オーソドックスな面接官配置の1つであろう。

図2は、全ての受験者に対して全ての質問を全ての面接官が採点を行う場合である。一番オーソドックスな面接官配置であろう。

図3は、各面接官が全ての質問ごとに1人の決められた受験者に対して採点を行う場合である。例えば、各自が質問を1つずつ行い、自分が行った質問に対してだけ採点を行うのであるが、面接としては非現実的な場面かもしれない。

図4は、各面接官に採点する質問と採点する受験者1人が決められていて、採点を行う場合である。

図5は、複数の面接室を設け、同じ面接官が同じ質問について採点を行う場合である。例えば、午前と午後（或いは、1日目と2日目）で二回にわけて同じ受験者に同じ質問を同じ面接官で行う場合が考えられるが、これも面接試験としては非現実的な場面である。

図6は、異なる面接官集団がいる複数の面接室を設け、その各面接室に受験生がまわっ

て、同じ質問について採点を行う場合である。例えば、午前と午後（或いは、1日目と2日目）で二回にわけて同じ受験者に同じ質問を違う面接官が行い評価する場合は考えられるが、これも図5同様、面接試験としては非現実的な場面である。

図7は、各面接室で異なる質問を行い、受験生が面接室をまわり、同じ面接官が採点を行う場合である。図5と図6と異なり、午前と午後（或いは、1日目と2日目）で面接官は変わらず、質問が異なる。しばしばみられるオーソドックスな面接官配置の1つであろう。

図8は、面接官が同じ複数の面接室があり、受験生は全ての面接室を周って同じ面接官から同じ質問を採点される。図5とほぼ同じであるが、面接官は質問したもののみ採点するところに違いがある。図5と同様、面接試験としては非現実的な場面である。

以下では、観点ごとに複数の質問を行い、観点の有用性を吟味したい場合に有効な面接官配置が続く。

図9は、各観点で複数の質問を行い、観点ごとに異なる面接官集団が採点を行う場合である。これも、図1と同様、例えば、アドミッション専任教員と学部教員が共同で面接したり、学部教員内で役割分担をしたりで、評価する質問項目が異なる場合には、この面接官配置が当てはまる。図1と異なるのは、観点ごとに複数の質問がされることである。これもオーソドックスな面接官配置の1つであろう。

図10は、各観点について面接官ごとに複数の質問を行い、自分が行った質問に対してのみ、採点を行う場合である。また、図11は、各観点について異なる面接官集団が設定され、自分が行った質問に対してのみ採点を行う場合である。

図10・11ともに、自分が行った質問に対してのみ採点をするという意味では、面接

試験としては非現実的な場面である。

図12は、面接室ごとに同じ面接官を配置し、全ての面接室を受験生がまわりながら、複数の観点で異なる面接官が採点する場合である。図9と左右対称になっているだけの実験配置である。図9と異なる点は、例えば、午前と午後（或いは、1日目と2日目）で、同じ面接官が同じ面接室に居るが、問う観点が面接官ごと異なり、問うた観点しか採点しないことである。

図13は、各観点で異なる質問項目があり、面接室ごとに受験生が割り振られて採点される場合である。また、図14は、各観点で異なる質問項目があり、面接室ごとに受験生が割り振られ、決まった観点の質問のみ採点される場合である。

図12・図13・図14は、面接室で1つの採点データしかないときに分析が可能な面接官配置である。

注

- 1) 実際に各要素は、交換可能である²⁾。従って、Brennan(2001)のTable. 9.2の実験計画において、面接場面で論理的にそぐわない場合、例えば、観点に質問項目がネストされている（具体的には、別の観点で同じ質問項目を行う）場合は、他の要素に変更してある。
- 2) 具体的に、木村・吉村(2010)の図4の実験計画は、本稿の図10において、iとhを入れ替えてある。

参考文献

- Brennan, R. L. (2001). *Generalizability Theory*, New York: Springer.
- 池田央 (1994). 「一般化可能性理論」『現代テスト理論』朝倉書店、28-50.
- 池田央(2007). 「一般化可能性理論」『統計データ科学事典』朝倉書店、638-9.
- 木村拓也・吉村宰(2010). 「AO入試における信頼性評価の研究——一般化可能性理論による検討」『大学入試研究ジャーナル』20、81-90.
- Shaverson, R. J. and Webb, N. M. (1991). *Generalizability Theory: A primer*, Newbury Park: Sage Publications.

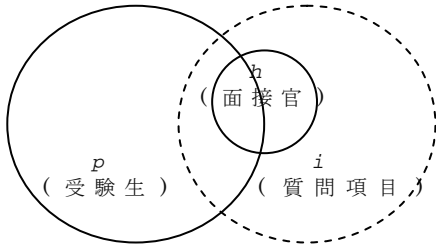


図 1. $p^* \times h^*$ デザイン

p	i_1					i_2				
	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}
1	X	X	X	X	X	X	X	X	X	X
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n_p	X	X	X	X	X	X	X	X	X	X

全ての受験者に対して質問ごとに異なる面接官集団が採点を行う場合

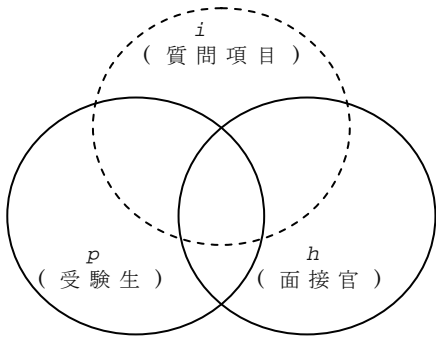


図 2. $p^* \times h^*$ デザイン

p	i_1					i_2				
	h_1	h_2	h_3	h_4	h_5	h_1	h_2	h_3	h_4	h_5
1	X	X	X	X	X	X	X	X	X	X
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n_p	X	X	X	X	X	X	X	X	X	X

全ての受験者に対して全ての質問を全ての面接官が採点を行う場合

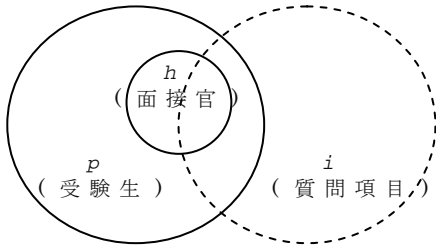


図 3. $h^* : p^*$ デザイン

p	i_1					i_2				
	h_1	h_2	h_3	h_4	h_5	h_1	h_2	h_3	h_4	h_5
1	X					X				
2		X					X			
3			X					X		
4				X					X	
5					X					X

各面接官が全ての質問ごとに1人の決められた受験者に対して採点を行う場合

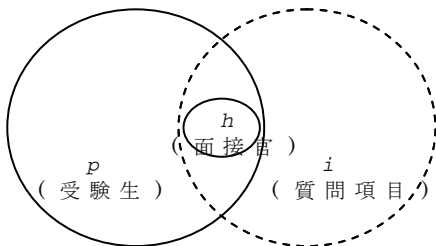


図 4. $h^* : p^*$ デザイン

p	i_1					i_2				
	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}
1	X					X				
2		X					X			
3			X					X		
4				X					X	
5					X					X

各面接官に採点する質問と担当する受験者1人が決められていて、採点を行う場合

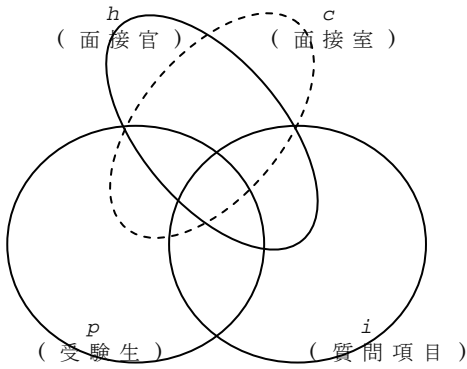


図 5. $p \times i \times h^\circ$ デザイン

p	c_1						c_2					
	h_1		h_2		h_3		h_1		h_2		h_3	
	i_1	i_2	i_1	i_2	i_1	i_2	i_1	i_2	i_1	i_2	i_1	i_2
1	X	X	X	X	X	X	X	X	X	X	X	X
•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•
n_p	X	X	X	X	X	X	X	X	X	X	X	X

複数の面接室を設け、同じ面接官が同じ質問について採点を行う場合

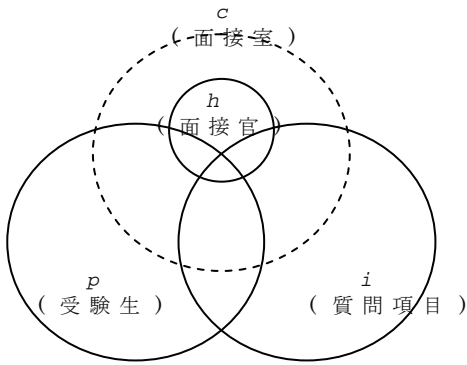


図 6. $p \times i \times h^\circ$ デザイン

p	c_1						c_2					
	h_1		h_2		h_3		h_4		h_5		h_6	
	i_1	i_2	i_1	i_2	i_1	i_2	i_1	i_2	i_1	i_2	i_1	i_2
1	X	X	X	X	X	X	X	X	X	X	X	X
•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•
n_p	X	X	X	X	X	X	X	X	X	X	X	X

異なる面接官集団がいる複数の面接室を設け、その面接室を受験生がまわって同じ質問について採点

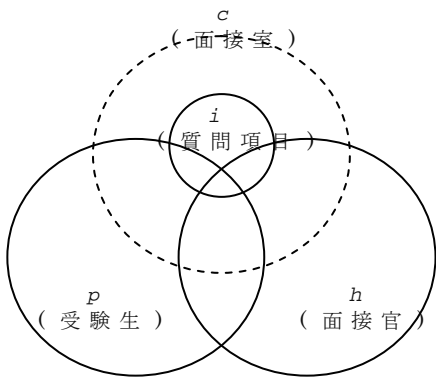


図 7. $p \times i \times h^\circ$ デザイン

p	c_1						c_2					
	i_1		i_2		i_3		i_4		i_5		i_6	
	h_1	h_2	h_1	h_2	h_1	h_2	h_1	h_2	h_1	h_2	h_1	h_2
1	X	X	X	X	X	X	X	X	X	X	X	X
•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•
n_p	X	X	X	X	X	X	X	X	X	X	X	X

各面接室で異なる質問を行い、受験生が面接室をまわり、同じ面接官が採点を行う場合

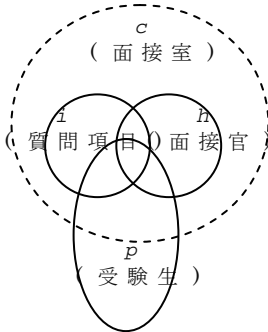


図 8. $p \times i \times h^\circ$ デザイン

p	c_1						c_2					
	h_1		h_2		h_3		h_1		h_2		h_3	
	i_1	i_2	i_3	i_4	i_5	i_6	i_1	i_2	i_3	i_4	i_5	i_6
1	X	X	X	X	X	X	X	X	X	X	X	X
•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•
n_p	X	X	X	X	X	X	X	X	X	X	X	X

面接官が同じ複数の面接室があり、受験者は全ての面接室を周って同じ面接官から同じ質問を採点

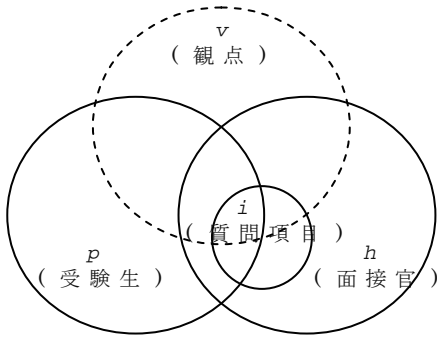


図 9. $p^* \times (i^* : h^*)$ デザイン

p	v_1						v_2					
	h_1		h_2		h_3		h_4		h_5		h_6	
	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8	i_9	i_{10}	i_{11}	i_{12}
1	X	X	X	X	X	X	X	X	X	X	X	X
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n_p	X	X	X	X	X	X	X	X	X	X	X	X

各観点で複数の質問を行い、観点ごとに異なる面接官集団が採点を行う場合

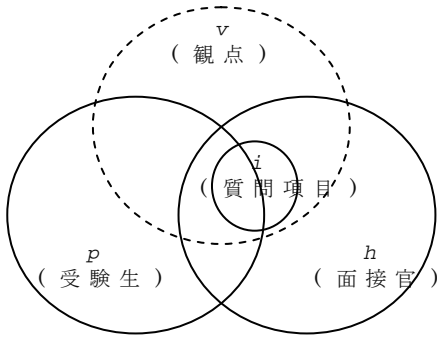


図 10. $p^* \times (i^* : h^*)$ デザイン

p	v_1						v_2					
	h_1		h_2		h_3		h_1		h_2		h_3	
	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8	i_9	i_{10}	i_{11}	i_{12}
1	X	X	X	X	X	X	X	X	X	X	X	X
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n_p	X	X	X	X	X	X	X	X	X	X	X	X

各観点について面接官ごとに複数の質問を行い、自分が行った質問に対してのみ、採点を行う場合

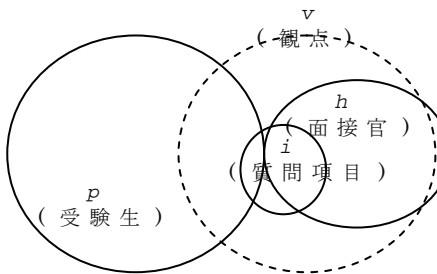


図 11. $p^* \times (i^* : h^*)$ デザイン

p	v_1						v_2					
	h_1		h_2		h_3		h_4		h_5		h_6	
	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8	i_9	i_{10}	i_{11}	i_{12}
1	X	X	X	X	X	X	X	X	X	X	X	X
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n_p	X	X	X	X	X	X	X	X	X	X	X	X

各観点について異なる面接官集団が設定され、自分が行った質問に対してのみ、採点を行う場合

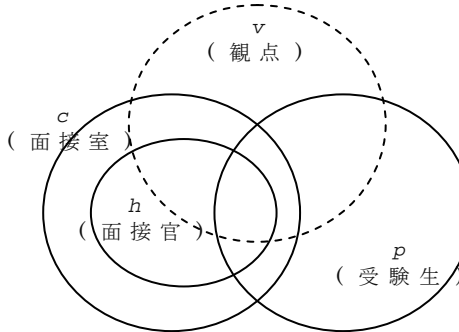


図 12. $(h^* : c^*) \times p^*$ デザイン

p	v_1						v_2					
	c_1		c_2		c_3		c_1		c_2		c_3	
	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}	h_{11}	h_{12}
1	X	X	X	X	X	X	X	X	X	X	X	X
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n_h	X	X	X	X	X	X	X	X	X	X	X	X

面接室ごとに同じ面接官を配置し、全ての面接室を受験生がまわりながら複数の観点で異なる面接官が採点する場合

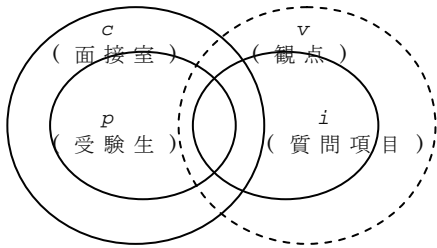


図 13. $(p^{\bullet}:c^{\bullet}) \times i^{\circ}$ デザイン

P	v_1						v_2					
	i_1			i_2			i_3			i_4		
	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3
1	X			X			X			X		
2		X			X			X			X	
3			X			X			X			X

各観点で異なる質問項目があり、面接室ごとに受験者が割り振られて採点される場合

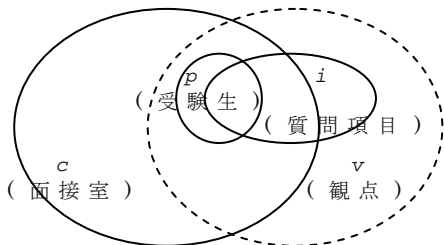


図 14. $(p^{\circ}:c^{\bullet}) \times i^{\circ}$ デザイン

P	v_1						v_2					
	i_1			i_2			i_3			i_4		
	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3
1	X			X								
2		X			X							
3			X			X						
4							X			X		
5								X			X	
6									X			X

各観点で異なる質問項目があり、面接室ごとに受験生が割り振られ、決まった観点の質問のみ採点する