

入学前教育におけるプレースメントテストの信頼性と AO・推薦入試合格者の学力推移

——2015～2018年度入学者のテストスコアの統計分析から——

菅 原 良*

Reliability of placement tests in pre-entrance education and academic achievement transition of those
who passed the admissions office examination and examination for selected candidates passer

— Based on statistical analysis of time series data for test score from FY 2015 to FY 2018 —

Ryo Sugawara *

1. 問題の所在

日本経済団体連合会(以下、経団連)は、2021年春入社以降の新卒者を対象とする就職・採用活動のルールを策定しないことを正式に決定した(日本経済団体連合会 2018)。その背景には、「アメリカの大学が行っているような入学後に所定の成績を取めた学生にだけ卒業を認める『出口管理』を徹底すること」(林 2018)が経済界から求められるようになったことが大きく関係していることが一因としてある。経済界からの要求は、経団連会長中西宏明による「これまで日本の若い人を見てきたが、グローバルな場での競争力やイノベティブな姿勢では、他国の若い人に見劣りする。自分で手を挙げ仕事を開拓する姿勢が必要。それは徒弟制度的な育成計画ではできない。経済界と教育界が議論し、大学の卒業証書が品質保証になる状況にしたい」(東洋経済新報社 2018)という、産学一体で大学教育の充実に取り組むべきだとする訴えによって端的に示される。

アメリカのAO入試をヒントに「日本で最初にAO入試を導入したのが1990年の慶応義塾大学湘南藤沢キャンパス(SFC)の2学部である」(林 2018)が、アメリカの大学では「出口管理」が徹底されているため、厳密な学力の精査を入学時に課す必要はないが、「日本は『入口管理』の国であるため、アメリカ方式そのまま導入してもうまく機能しないため、高等学校での学習履歴だけでなく詳細な書類選考と時間をかけた丁寧な面接等を組み合わせて受験者の能力・適性や学習に対する意欲・目的意識等を総合的に判定する方式に改変して導入された」(林 2018)。日本のAO入試が、「アメリカのAO入試が、志願者に対して履修科目や成績といった高等学校での学習履歴を確認し、入学後の修学に耐える準備を行ってきているかをチェックすることに主眼を置いた入試である」(林 2018)こととは大きく異なる。

しかし日本のAO入試では、大学(或いは同じ大学でも学部・学科)によっては、学力を測ることが行わないことも多いため、「AO・推薦入試合格者の入学前段階における学力(基礎的・基本的な「知識や技能」、文部科学省)は、高等学校から提出される資料でしか把握することができない」(菅原 2018)ことも多い。また、高等学校により資料の信頼性が異なることから、これらの資料によって生徒の学力を入学前段階において、同一尺度上で把握し比較することは困難である(菅原 2017)。このような諸課題に対処することを目的として、明星教育センターでは、全学部のAO・推薦入試(指定校、公募制、卒業生子女、スポーツ文化、明星高校特別)合格者を対象として、共通の入学前教育プログラムを実施しており、2015年度入学生から入学前教育の一部を、eラーニングによる自宅学習を課すことにより、入学前教育プログラムにおいて実施するプ

* 明星教育センター特任教授

レースメントテスト（以下、プレテスト）のスコアで学力を可視化している。それによって、AO・推薦入試合格者の学力を入学前段階において同一尺度上で把握し、経年比較することができる仕組みを整えている。しかし、2015年度から導入したプレテストが、合格者の学力を的確に測れているか（妥当性）、安定した信頼のおけるものであることが保証されたものであるか（信頼性）、についての議論は行われてこなかった。

本稿ではこれらの問題意識により、最初に入学前教育で合格者に受験させているプレテストの妥当性と信頼性について検討する。次に各学科の合格者のプレテストスコアの統計分析から、AO・推薦入試合格者の学力推移を客観的に把握し、AO・推薦入試の妥当性と入学後の大学教育における質保証に関して議論する。

2. プレースメントテストの妥当性と信頼性

明星教育センターでは、2014年度（2015年度入学生）以降、入学前教育の一環としてAO・推薦入試合格者を対象としてeラーニングを受講させており、その難易度を決定するために受験させたプレテスト（英語・数学・国語・理科）約4,000件（4,000人分）のスコアを保存している。

ここで重要なのがプレテストの妥当性 (validity)¹と信頼性 (reliability)²である。ここでは、受験させているテストが、学力を的確に測れているか（妥当性）、安定した信頼のおけるものであることが保証されたものであるか（信頼性）、について検討する。テストの妥当性については、内容的妥当性（専門家によって判断される項目の適切性）に基づいている。プレテストの妥当性については、テスト作成に通じた大学教員および高等学校教員といった専門家によって作成されていることから一定程度の水準は担保されていると考える。また、信頼性については、テストの信頼性係数（クロンバックの α 係数）を算出することによって測定することができる。 m は質問の項目数、 σ_i は各質問項目の分散、 σ_x は各質問項目を合計した尺度スコアの分散としたとき、下式から α 係数が求められる。

$$\alpha = \frac{m}{m-1} \left(1 - \frac{\sum_{i=1}^m \sigma_i^2}{\sigma_x^2} \right)$$

英語プレテストの α 係数は、2015年度が .785 と高かったが、2016年度から2018年度ではいずれも .700 を下回っている。国語では、2015年度は .560 と信頼性は低かったが、2017年度は .691、2018年度は .744 となった（2016年度はデータ欠損のため、算出することができなかった）。数学では、2015年度には .799 と高かったが、2016年度以降は .700 を下回っている。一般に α 係数は .70 から .80 以上あるとテストの信頼性は高いとされており、今回の検討対象となったプレテストでは、2015年の英語および数学、2018年の国語が .700 を上回っているが、それ以外はいずれも .700 を下回っており、プレテストの信頼性は総じてそれほど高いとは云えなかった。

しかし、本稿で扱うプレテストは、入学試験などのような厳密に信頼性が問われる性質のものではないことから、信頼性は許容の範囲としても問題ないと思われる。ただし、テストの妥当性と信頼性には、クロンバック (Cronbach) が「帯域幅と忠実度のジレンマ (bandwidth-fidelity dilemma)」と呼んだように、妥当性を高めようとするれば信頼性が犠牲になり、信頼性を高めようとするれば妥当性が犠牲になってしまうという性質があるため、現実的にはこの双方を同時に満たすことは困難であるとされていることに留意を要する（産業能率大学総合研究所 2017）。

表1 プレテストのクロンバックの α 係数

入学年度	英語	国語	数学
2015	<u>.785</u>	.560	<u>.799</u>
2016	.633	-	.653
2017	.677	.691	.667
2018	.695	<u>.744</u>	.678

3. プレースメントテストの通過率

受験させたテストスコアを経年で比較することができるように、各年度のテストのうち、いくつかのテスト項目を共通で出題している。これらのテスト項目の通過率を比較することによって AO・推薦入試合格者の学力を経年比較することができる。本稿では、英語、国語、数学について検討する。

英語では、通過率が極端に低い項目は En 20 および En 24 であった。対応するテスト項目の内容は、「いろいろな SVOC」と「I was, as if の表限」の2問であった(表2)。国語では、2017年度の Nl 04 の通過率が極端に低かった(表3)。数学では、Mt 05 で高い通過率となった。対応するテスト項目の内容は、「三角関数、三角方程式、加法定理とその応用」であった。一方、通過率が .100 を下回るテスト項目はなかった(表4)。また、年度によって通過率が大きく異なる項目はみられなかった。

表2 通過率(英語)の推移(年度比較)

項目	2015	2016	2017	2018	項目	2015	2016	2017	2018
En01	.707	.716	.709	.724	En13	.290	.256	.294	.264
En02	.830	.838	.820	.873	En14	.268	.223	.267	.282
En03	.479	.454	.467	.493	En15	.164	.124	.143	.151
En04	.581	.600	.597	.579	En16	.378	.345	.327	.364
En05	.860	.831	.809	.860	En17	.430	.440	.426	.478
En06	.422	.473	.494	.479	En18	.293	.299	.282	.316
En07	.304	.270	.290	.287	En19	.392	.351	.380	.384
En08	.652	.701	.673	.718	<u>En20</u>	<u>.055</u>	<u>.056</u>	<u>.049</u>	<u>.143</u>
En09	.718	.685	.710	.698	En21	.129	.143	.154	.143
En10	.282	.309	.285	.326	En22	.526	.484	.466	.505
En11	.548	.575	.597	.588	En23	.268	.305	.275	.322
En12	.622	.657	.652	.680	<u>En24</u>	<u>.071</u>	<u>.067</u>	<u>.076</u>	<u>.099</u>

表3 通過率(国語)の推移(年度比較)

項目	2017	2018	項目	2017	2018	項目	2017	2018
Nl01	.106	.162	Nl10	.573	.518	Nl19	.916	.921
Nl02	.553	.755	Nl11	.247	.490	Nl20	.963	.949
Nl03	.278	.304	Nl12	.976	.987	Nl21	.770	.700
Nl04	<u>.038</u>	.168	Nl13	.910	.891	Nl22	.720	.742
Nl05	.569	.526	Nl14	.409	.441	Nl23	.850	.833
Nl06	.721	.722	Nl15	.787	.808	Nl24	.453	.480
Nl07	.450	.490	Nl16	.745	.785	Nl25	.820	.811
Nl08	.516	.335	Nl17	.748	.591	Nl26	.968	.866
Nl09	.561	.535	Nl18	.924	.908			

※ 2015、2016年度は保存データが破損したため算出することができなかった。

表4 通過率(数学)の推移(年度比較)

	2015	2016	2017	2018		2015	2016	2017	2018
Mt01	.428	.553	.545	.574	Mt05	.842	.891	.908	.901
Mt02	.576	.680	.687	.667	Mt06	.502	.493	.532	.525
Mt03	.448	.548	.535	.623	Mt07	.606	.654	.675	.677
Mt04	.717	.771	.741	.778	Mt08	.279	.311	.333	.323

4. 平均スコアの推移

各年度の受験者の群毎の能力値を比較するために、2015年度以降実施してきたテストによって取得したスコアの5分位図を作成した。

英語では、2015年度の下位群(25.34)よりも2016年度の下位群(23.41)と2017年度の下位群(25.11)の平均点が低くなったが、それ以外の年度では前年度よりも平均点が低くなった群はなかった。2015年度と比較して2018年度の平均点の伸びは、中位群、中上位群、上位群でより高くなっており、特に上位群の平均点の伸びが著しい。得点上位層の学力の伸びが大きくなっている一方、下位群の底上げはほとんどみられなかった(表5、図1)。

表5 プレテスト(英語)の平均スコアの5分位(2015-2018)

	2015		2016		2017		2018		2018-2015
	\bar{x}	N	\bar{x}	N	\bar{x}	N	\bar{x}	N	\bar{x}
下位群	25.34	73	23.41	197	25.11	192	25.96	212	.62
中下位群	33.75	73	34.03	197	35.21	192	36.88	211	3.13
中位群	39.53	73	41.14	197	42.28	192	44.34	211	4.81
中上位群	47.81	73	48.35	197	50.52	192	52.23	211	4.42
上位群	61.62	73	62.58	197	66.53	192	67.96	211	6.34

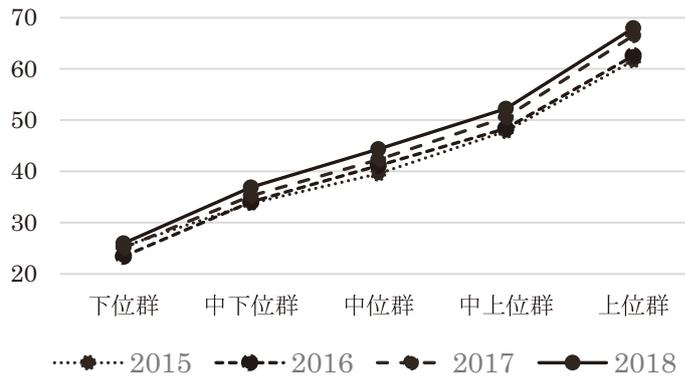


図1 プレテスト(英語)の平均スコアの5分位(2015-2018)

国語では、2017年度と比較して2018年度の平均点の伸びは、下位群で大きく低下し、中下位群で少し低下しているが、中位群、中上位群、上位群ではわずかではあるが伸張している。全体として学力の伸びは頭打ちの傾向にあることが推測される(表6、図2)。

表6 プレテスト(国語)の平均スコアの5分位(2015-2018)

	2015		2016		2017		2018		2018-2015
	\bar{x}	N	\bar{x}	N	\bar{x}	N	\bar{x}	N	\bar{x}
下位群	-	-	-	-	45.99	192	42.01	211	-3.98
中下位群	-	-	-	-	62.96	192	62.48	210	-.48
中位群	-	-	-	-	71.29	192	72.07	210	.78
中上位群	-	-	-	-	77.87	191	78.80	210	.93
上位群	-	-	-	-	86.19	191	87.13	210	.94

注1 2015、2016年度は保存データが破損したため算出することができなかったことから、平均点の伸びは、2018年度から2017年度のスコアを差し引いたものを参考として掲出した。

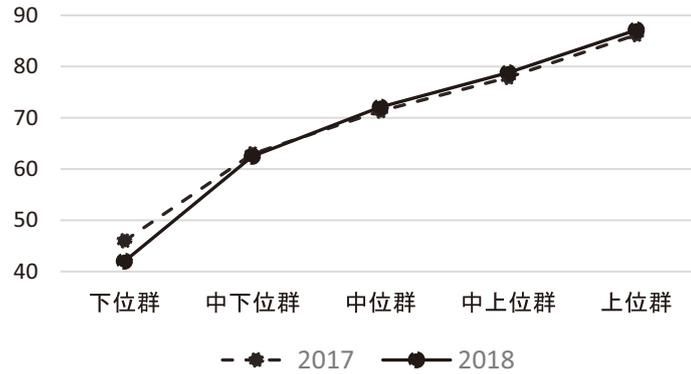


図2 プレテスト (国語) 平均スコアの5分位 (2015-2018)

数学では、2015年度と比較して2018年度の平均点はどの群でも大きく伸長している。特に中下位群、中位群、中上位群、上位群の伸びが顕著である(表7、図3)。

表7 プレテスト (数学) の平均スコアの5分位 (2015-2018)

	2015		2016		2017		2018		2018-2015
	\bar{x}	N	\bar{x}	N	\bar{x}	N	\bar{x}	N	\bar{x}
下位群	21.83	60	24.41	179	25.26	175	28.75	192	6.92
中下位群	37.67	60	46.54	179	46.06	175	54.32	192	16.65
中位群	50.17	59	59.04	178	60.69	175	68.02	192	17.85
中上位群	64.24	59	72.64	178	74.48	174	80.00	192	15.76
上位群	81.44	59	88.09	178	89.60	174	92.93	191	11.49

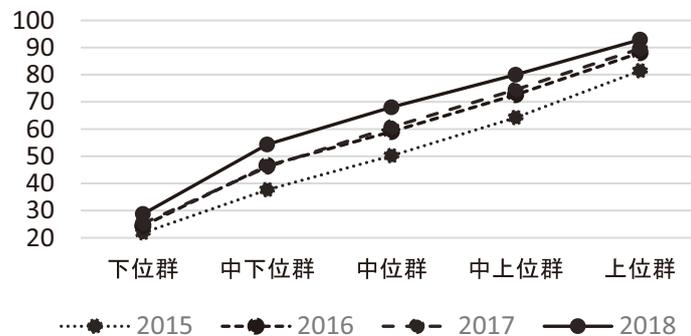


図3 プレテスト (数学) の平均スコアの5分位 (2015-2018)

5. 5分位図からみた AO・推薦入試の学科別合格者の特徴

学科毎の5分位を検討することによって、各学科の AO・推薦入試合格者が5分位のどの群に多く存在するかがわかる。本稿で検討した英語、国語、数学のすべての科目で、2015年度から2018年度を通じて、教育学科が上位群の生徒を合格させていることが特徴的である(図4、図5、図6、図7、図8、図9、図10、図11、図12、図13)。

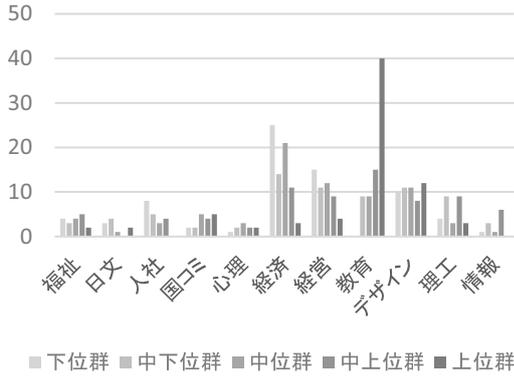


図4 AO・推薦入試の学科別合格者の5分位 (英語) 2015年度

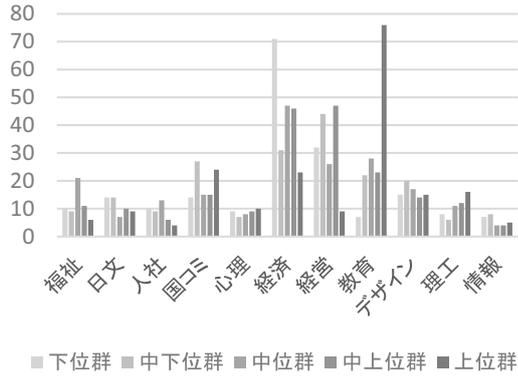


図5 AO・推薦入試の学科別合格者の5分位 (英語) 2016年度

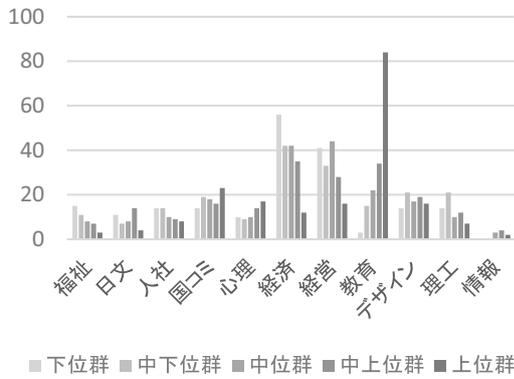


図6 AO・推薦入試の学科別合格者の5分位 (英語) 2017年度

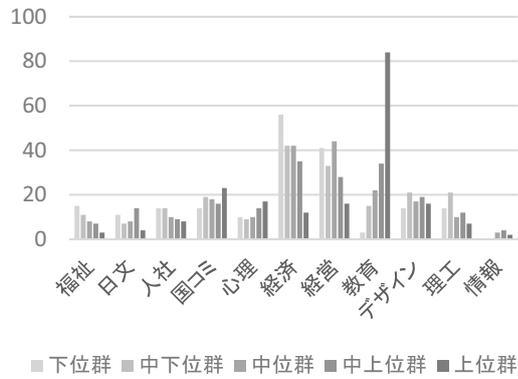


図7 AO・推薦入試の学科別合格者の5分位 (英語) 2018年度

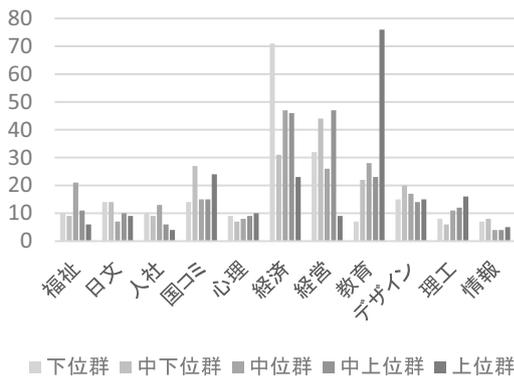


図8 AO・推薦入試の学科別合格者の5分位 (国語) 2017年度

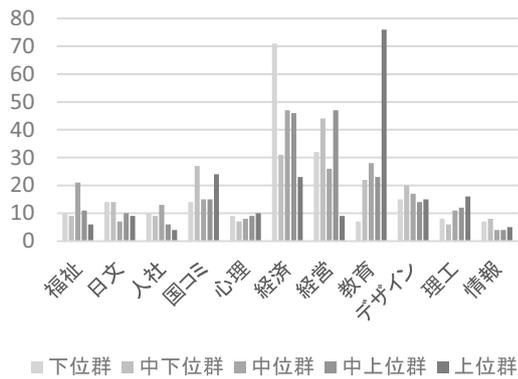


図9 AO・推薦入試の学科別合格者の5分位 (国語) 2018年度

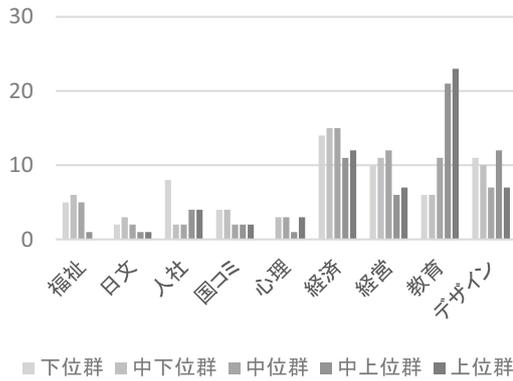


図 10 AO・推薦入試の学科別合格者の 5 分位 (数学) 2015 年度

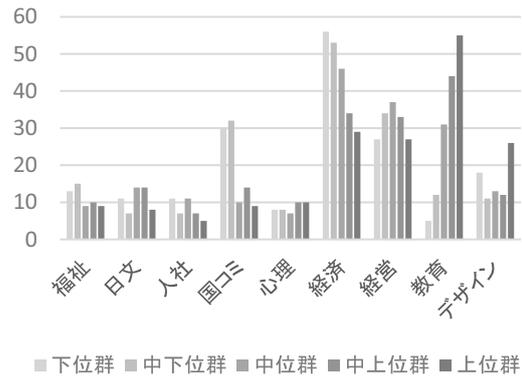


図 11 AO・推薦入試の学科別合格者の 5 分位 (数学) 2016 年度

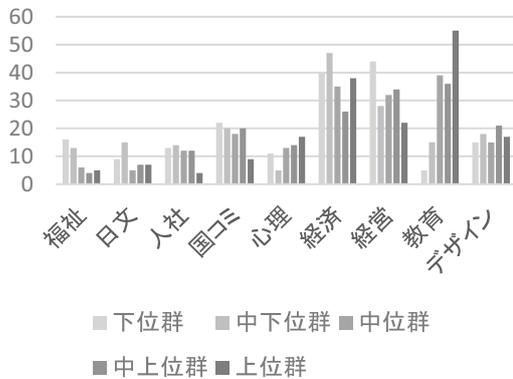


図 12 AO・推薦入試の学科別合格者の 5 分位 (数学) 2017 年度

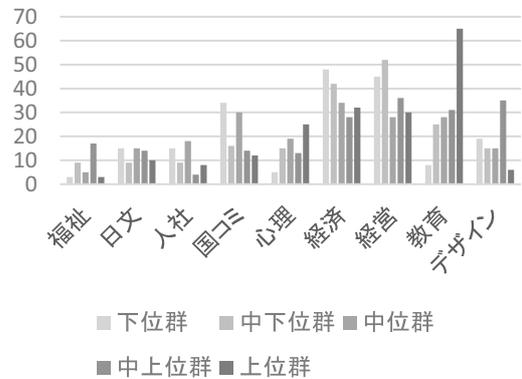


図 13 AO・推薦入試の学科別合格者の 5 分位 (数学) 2018 年度

6. 考察

最初に入学前教育で合格者に受験させているプレテストの妥当性と信頼性について検討した。その結果、テストの妥当性および信頼性について、一定程度の水準は担保されていることがわかった。次に各学科の合格者のプレテストスコアの統計分析から、AO・推薦入試合格者の学力推移について検討した。5分位で検討した結果、教育学科が上位群の生徒を合格させていることがわかった。その他にも学科によって合格者の学力について特徴があることがわかった。本研究で得られた知見は、明星大学における AO・推薦入試のあり方を考えていくうえで重要な示唆を提供することができたのではないと思われる。

注

- 1 その検査が測定しようとしているものをどれくらい的確に測定できているか、ということ。妥当性には、内容的妥当性(専門家によって判断される項目の適切性)、基準関連妥当性(尺度と関連性が強いと考えられる外的基準との相関によって評価される妥当性)、構成概念妥当性(尺度スコアの高低が、構成概念の強弱を支持しているかどうかを表す)、などがある。しかし現在では、内容的妥当性と基準関連妥当性が構成概念妥当性に吸収され、「妥当性はひとつの統合体である」という考え方が主流になりつつある。

- 2 仮に同じ条件の下で同じ検査を受けた場合、同じような結果が出る(結果が一貫し、安定している)ということ。信頼性には、再テスト法、折半法、内部一貫性による方法がある。

文献

- 林篤裕(2018)アドミッション・オフィスの機能と役割－多面的・総合的評価を実現するために－, 名古屋高等教育研究, 第18号, pp.39-53
- 日本経済団体連合会(2018)定例記者会見における中西会長発言要旨 <http://www.keidanren.or.jp/speech/kaiken/2018/1009.html> (2019.11.11 閲覧)
- 日本経済新聞社(2018)新卒一括採用、転職に 経団連が就活ルール廃止発表 <https://www.nikkei.com/article/DGXMZO36281670Z01C18A0MM8000/> (2019.11.11 閲覧)
- 産業能率大学総合研究所(2017)人材開発活動に必要なテストの作成方法と考え方・すすめ方 <https://www.hj.sanno.ac.jp/cp/feature/201703/09-06.html> (2019.11.11 閲覧)
- 菅原良(2017)AO・推薦入試合格者の学力推移と学習傾向－入学前教育におけるプレースメントテストおよび修了テストの時系列データの統計分析から－, 明星－明星大学明星教育センター研究紀要, 第8号, pp.17-25
- 菅原良(2018)2018年度入学生におけるAO・推薦入試合格者の学力推移と学習傾向－入学前教育eラーニングのプレースメントテストおよび修了テストにおける時系列データの統計分析から－, 明星－明星大学明星教育センター研究紀要, 第9号, pp.49-59
- 東洋経済新報社(2018)特集 就活ルール廃止の衝撃 経団連／新経連／就職問題懇談会の幹部が語る就活ルールのあるべき姿とは? <https://premium.toyokeizai.net/articles/-/19079> (2019.11.11 閲覧)