Results of a Data Visualization Competition to Report Item Performance Statistics

Dr. Stephanie Evergreen
Alex Kassidis
Dr. Manny Straehle

ATP Innovations in Testing 2014
Scottsdale, AZ

Agenda

History

Stephanie

Evergreen

Competition

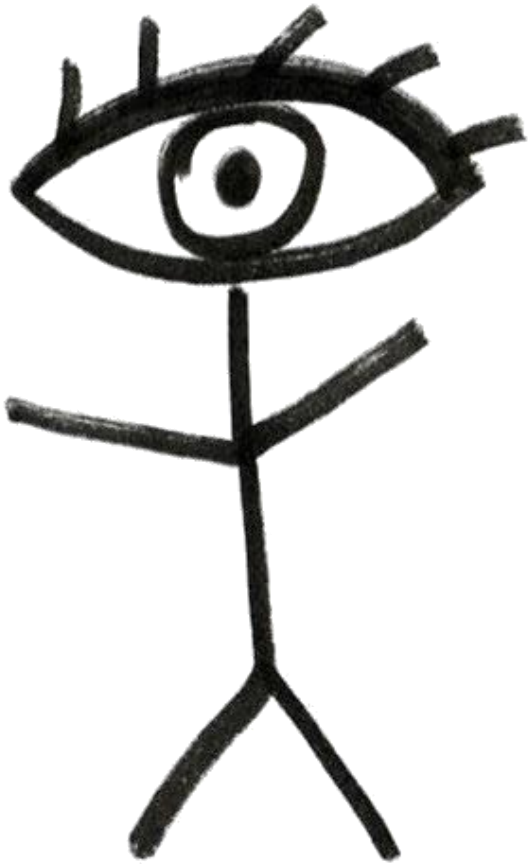
Alex

Kassidis

Results

Manny

Straehle

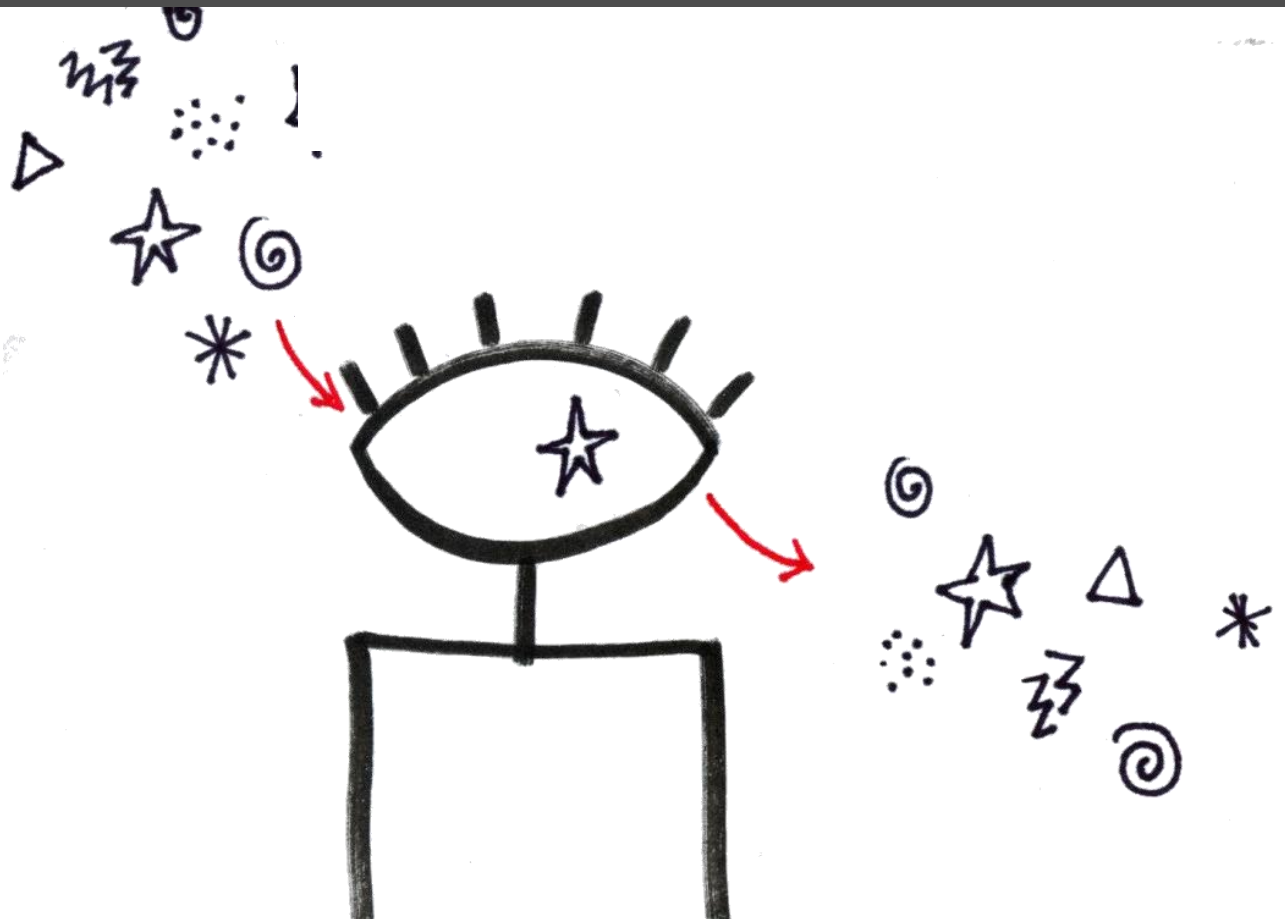


Pictorial Superiority Effect

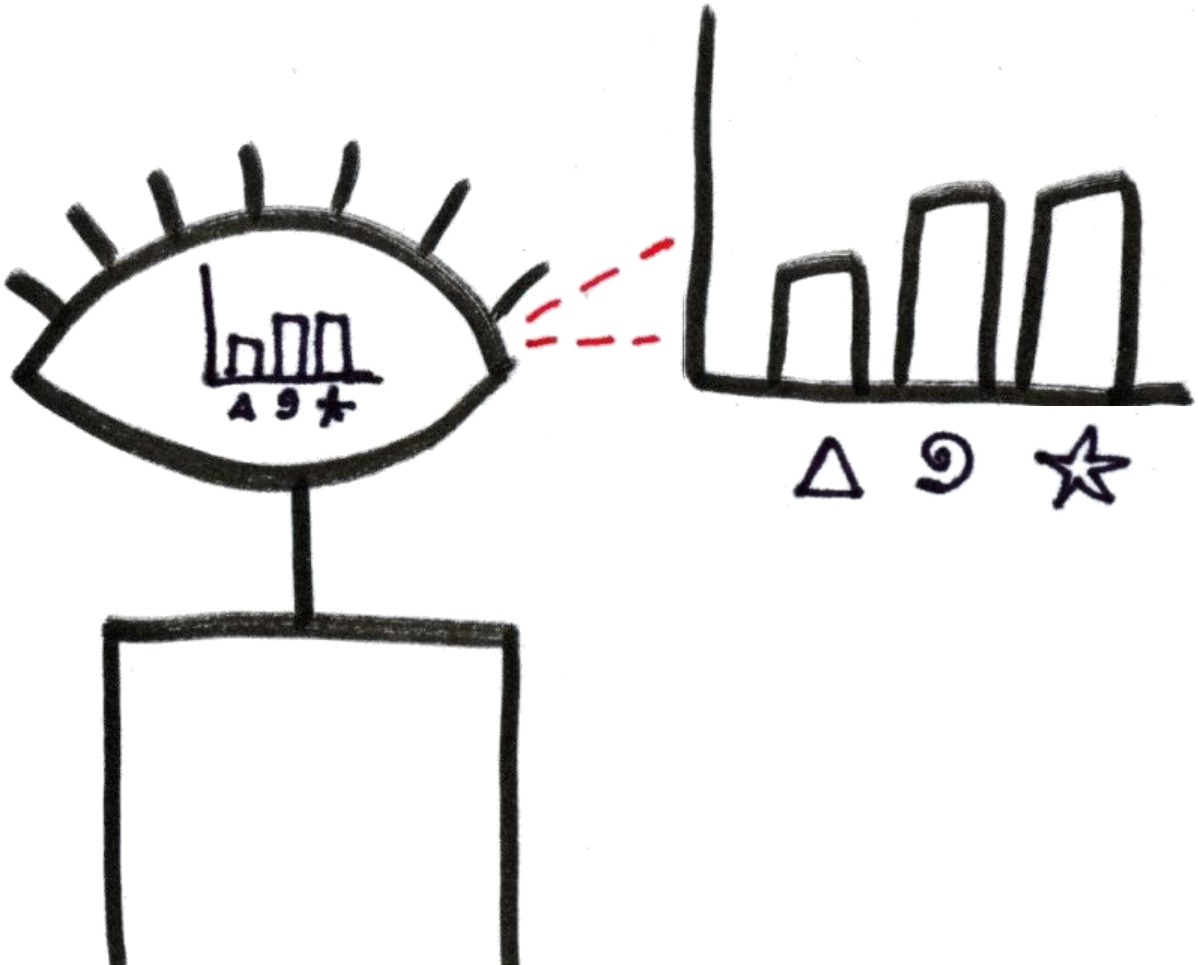
tiger?

oops?

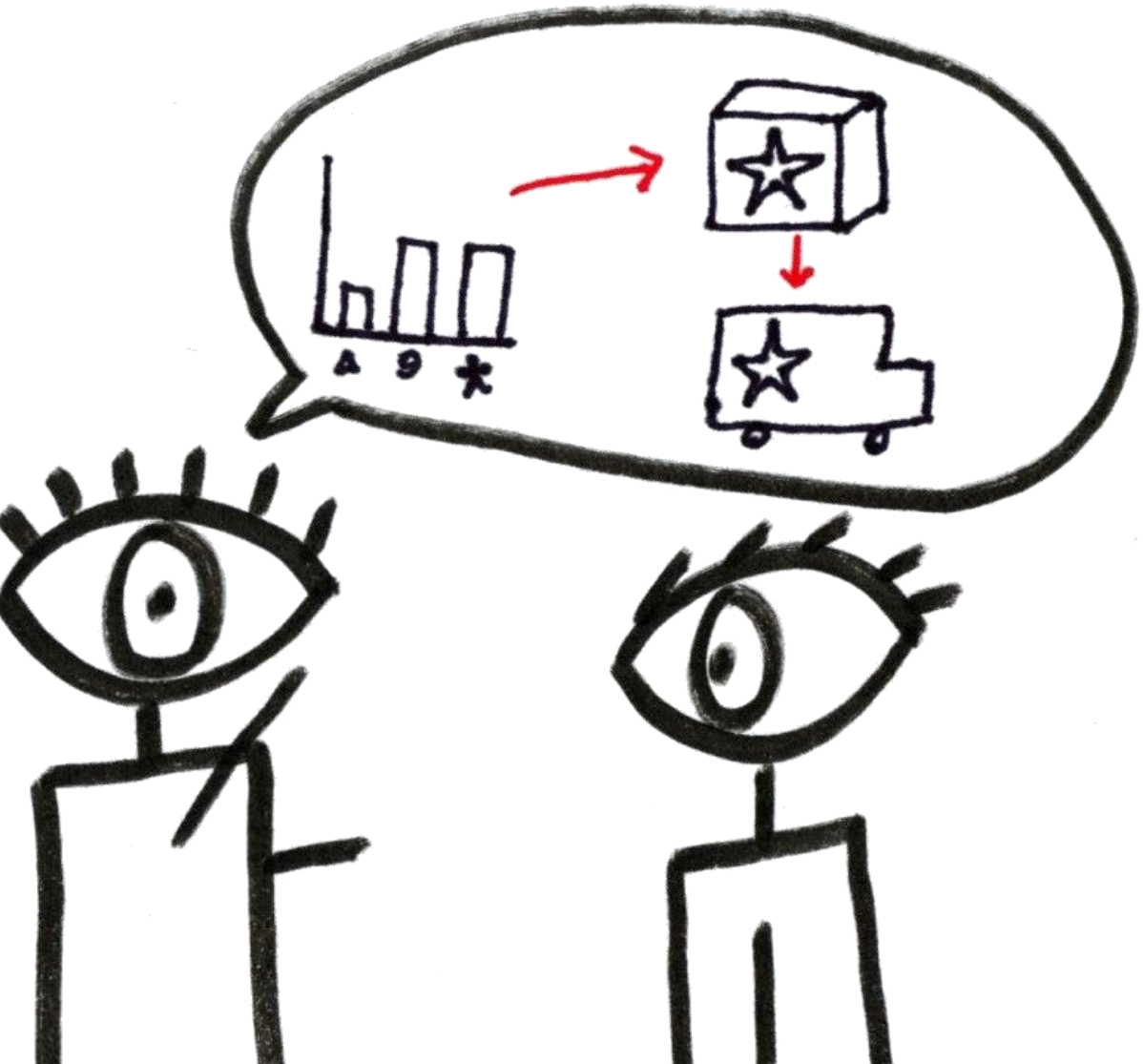
working memory



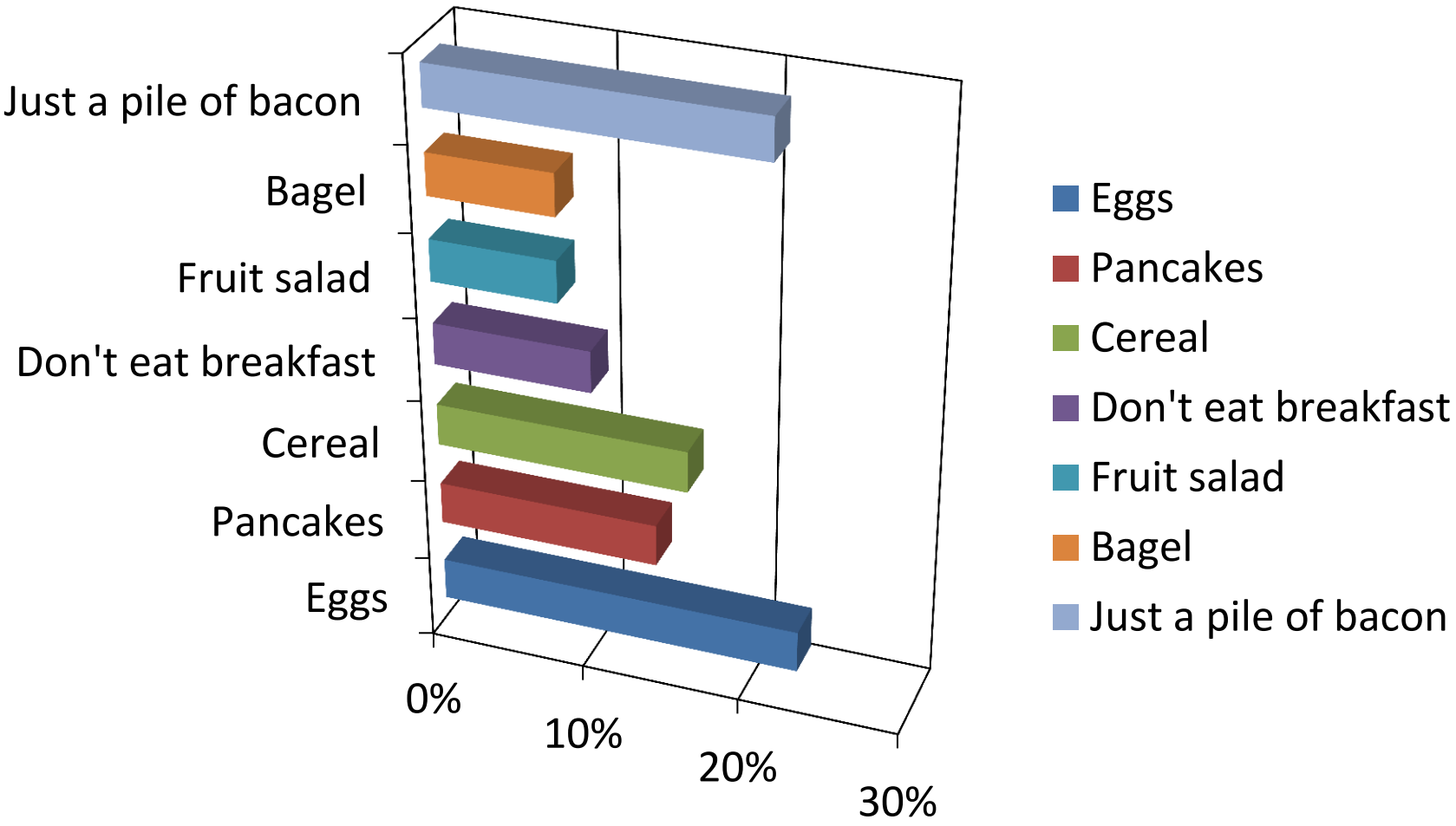
organize & emphasize



long term memory

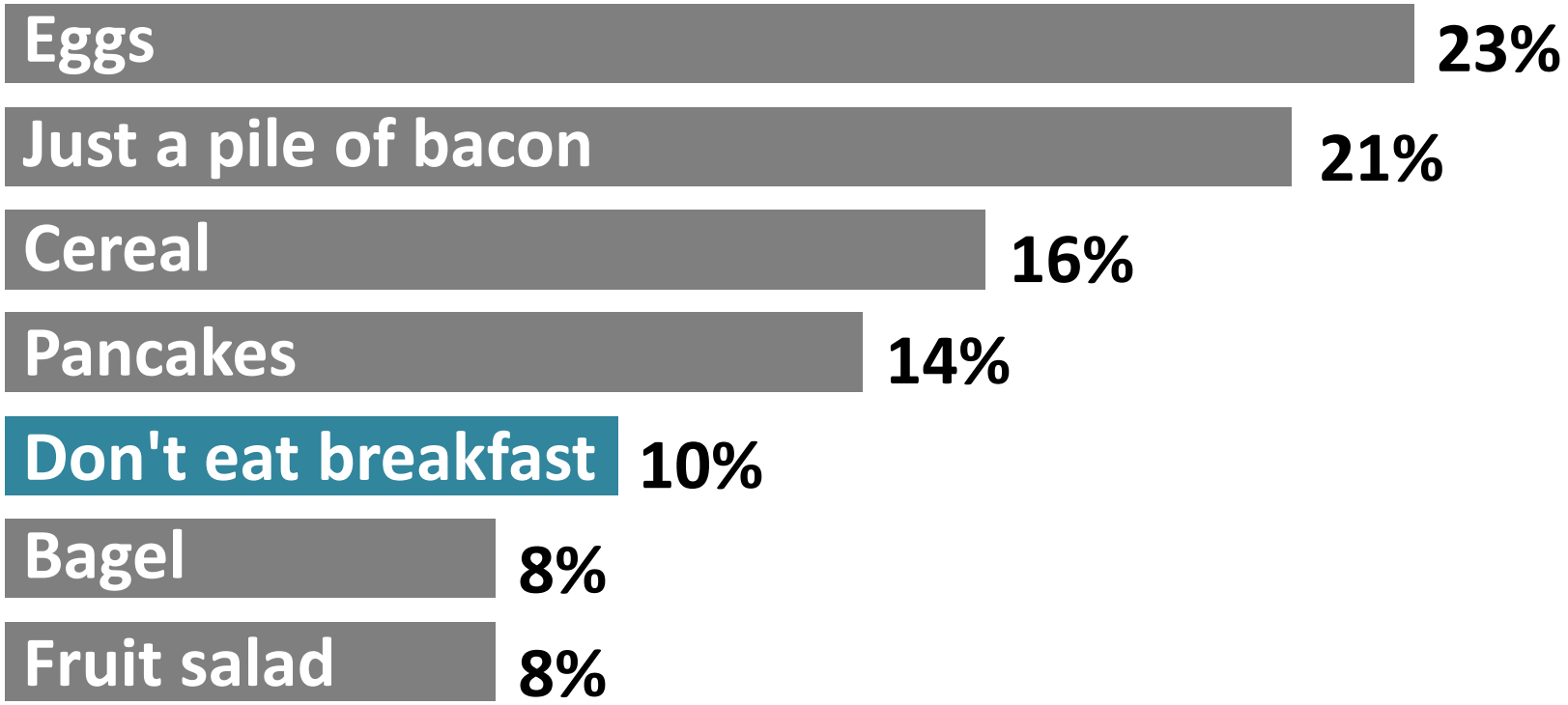


Tester Breakfast Preferences

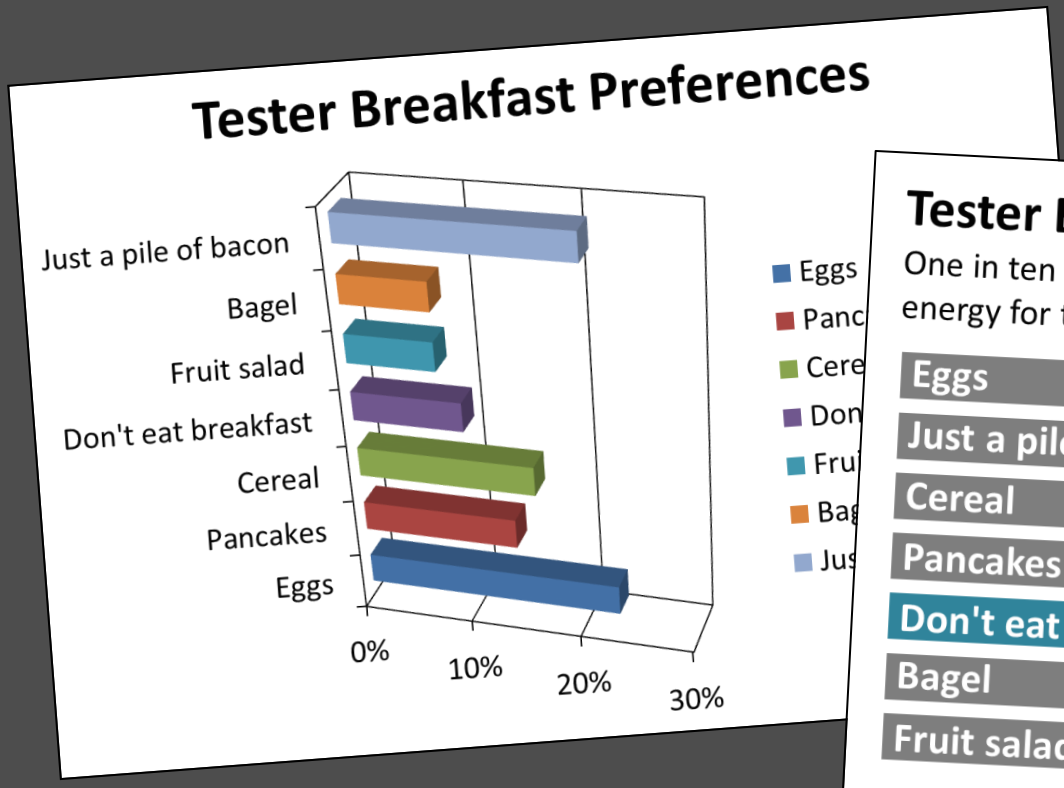


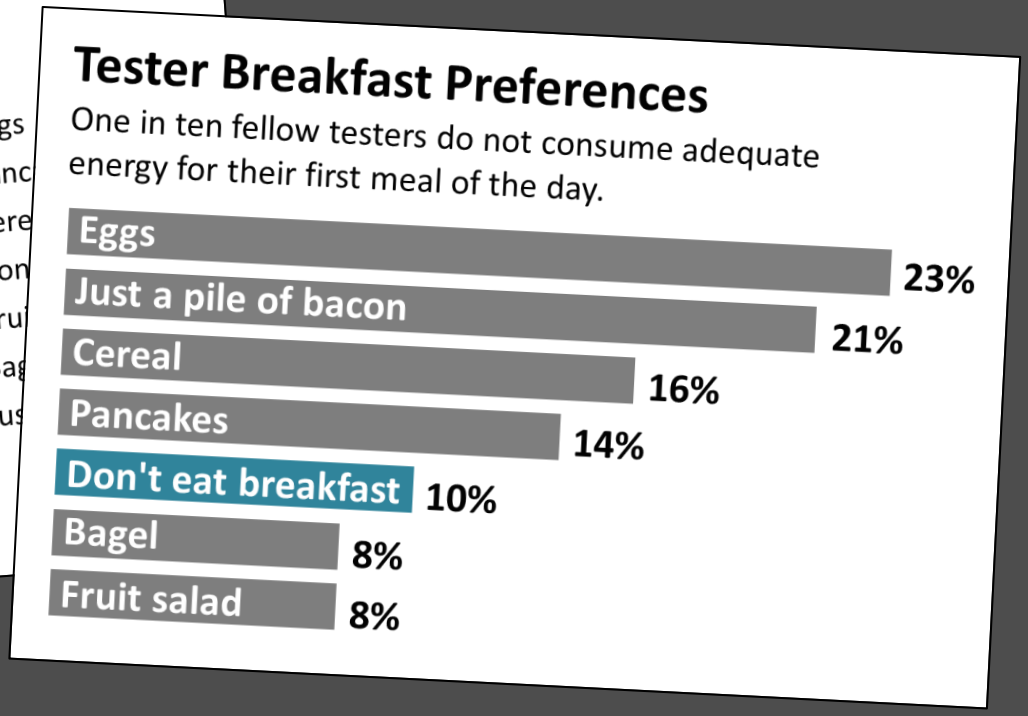
Tester Breakfast Preferences

One in ten fellow testers do not consume adequate energy for their first meal of the day.

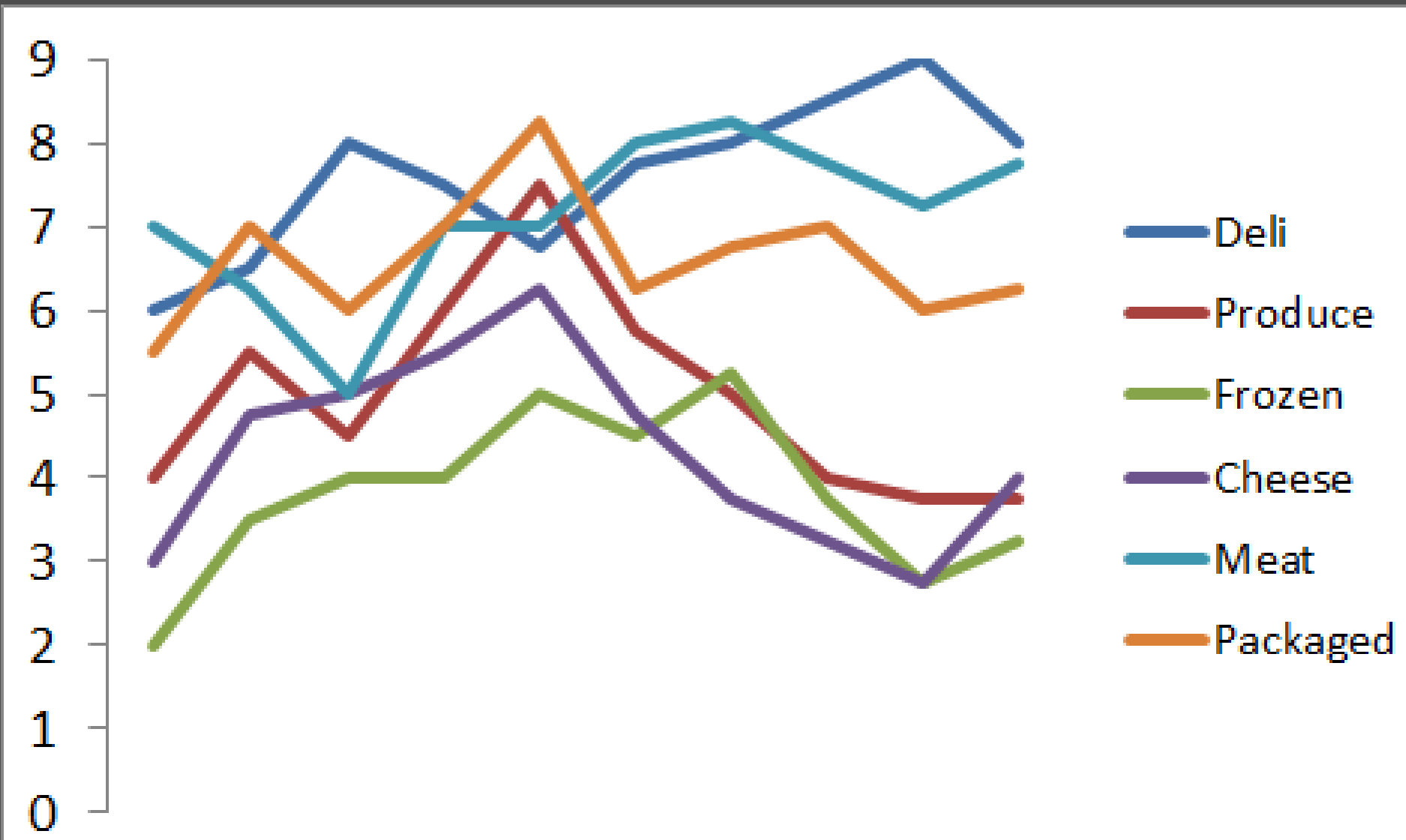


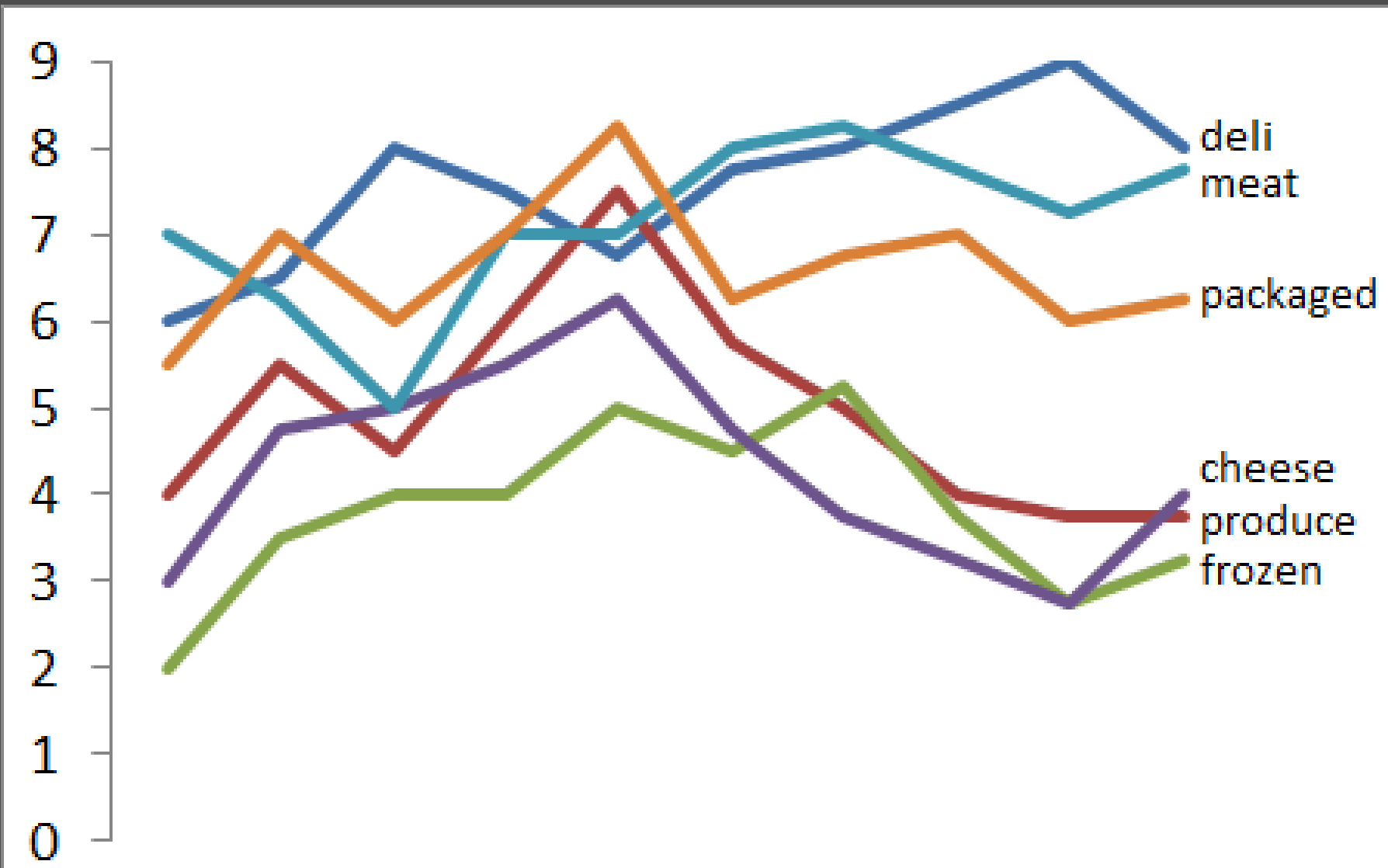
before & after





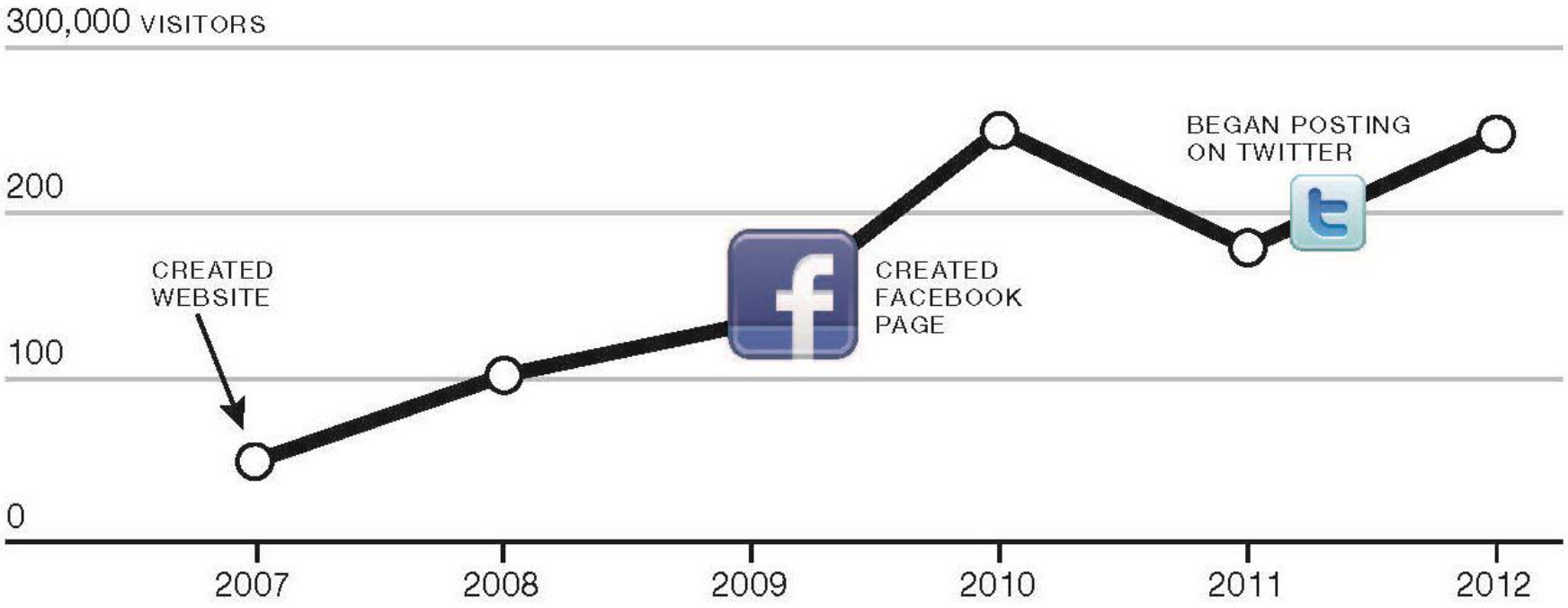
ATP 2014
Innovations in Testing
Scottsdale, AZ


9
8
7
6
5
4
3
2
1
0
Deli
Produce
Frozen
Cheese
Meat
Packaged


9
8
7
6
5
4
3
2
1
0
deli
meat
packaged
cheese
produce
frozen

Using social media websites to attract customers

Visits increased after launching social media, then fell a bit with the national trend.



ATP Innovations in Testing 2014 Scottsdale, AZ

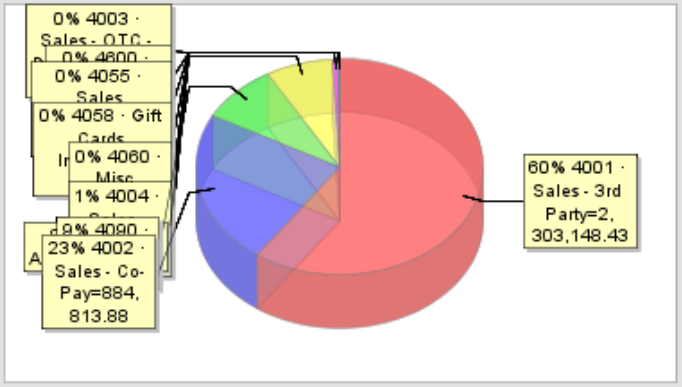
Class [-]

All Classes

exceldashboard.org

Income Distribution [-]

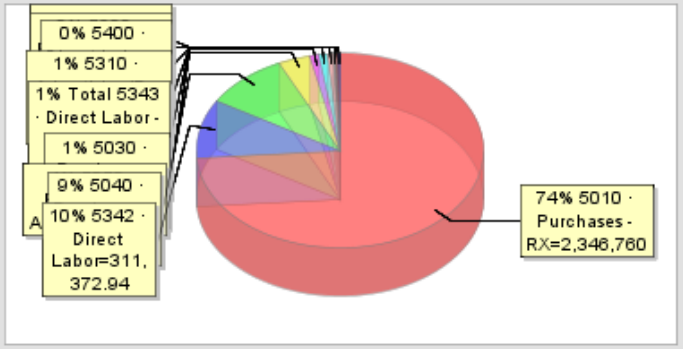
Income Distribution



4001 · Sales - 3rd Party · 4002 · Sales - Co-Pay · 4090 · HM
4095 · AD · 4004 · Sales - OTC · 4060 · Misc. Income
4058 · Gift Cards Income · 4055 · Sales Discounts
4600 · Merchandise Sales
4003 · Sales - OTC - Diabetes/Breathin

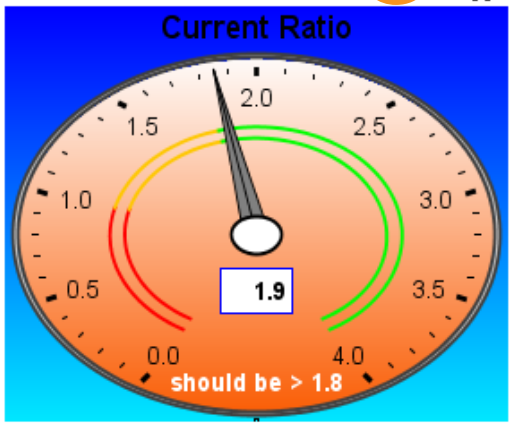
< COGS Distribution > [-]

COGS Distribution



5010 · Purchases - RX · 5342 · Direct Labor
5040 · Purchases - HM · 5050 · Purchases - AD
5030 · Purchases - Other RX
Total 5343 · Direct Labor - Payroll Taxes
5310 · Advertising/Promotion · 5400 · Direct Labor - Insurance
5200 · Freight/Shipping · Total 5700 · Automobile Exp.

Current Ratio [-]

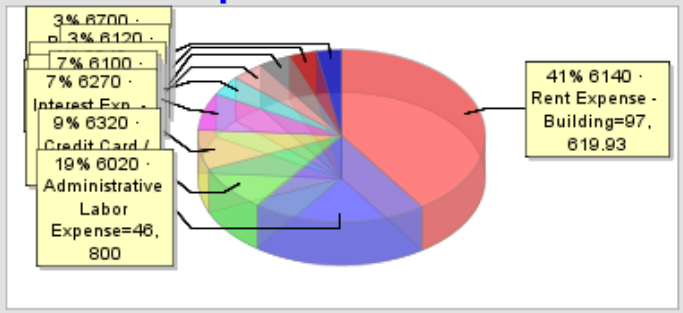


AR / AP Ratio [-]

Debt to Equity Ratio [-]

Expense Distribution [-]

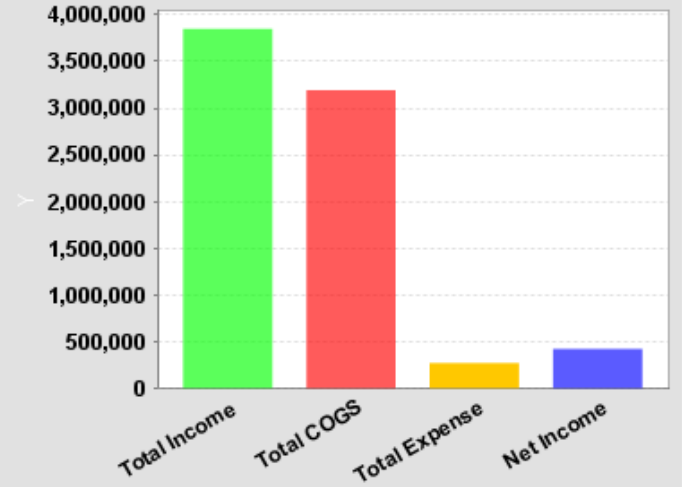
Expense Distribution



6140 · Rent Expense - Building
6020 · Administrative Labor Expense
6320 · Credit Card / Merch. Fees
6270 · Interest Exp. - Line of Credit · 6100 · Office Supply
6080 · Accounting & Legal · 6260 · Miscellaneous Exp.
6040 · Admin Labor - Insurance · 6120 · Travel Expense

Profit and Loss [-]

Profit and Loss



perceptualedge.com

executive dashboard

this data is as of november 15, 2005

Help

key figures

actual | target | bad | satisfactory | good

	actuals nov 04 - oct 05	current month	% of previous year	year to date	year to date % target
$ Revenues in mio		9,38	67,1	162,68	96,2
$ Expenses in mio		6,61	71,2	114,07	96,4
$ Profits in mio		2,76	58,9	61,11	120,4
Market Share in %		47,0	111,9	46,0	100,0
Flights		231	54,4	4.602	96,4
Passengers in K		32,4	61,1	595,2	95,2
Miles in K		163	51,0	3.513	103,9
Passenger Miles in mio		14,6	57,1	312,1	92,3
Cancelled Flights		7	58,3	104	80,0
Late Arrivals		48	58,5	714	178,9
Number of Minutes Late		1.342	59,3	15.363	187,8
$ Fuel Costs in mio		0,69	72,2	12,76	125,8
Average Customer Satisfaction Scores (out of 5)		4,5	99,6	4,5	94,0
Flight Utilization (Passengers / Capacity) in %		94,0	109,3	94,0	101,1
Revenue / Passenger Mile		0,6	117,4	0,6	128,8
Profit / Mile		17,0	115,4	17,0	113,3
Revenue / Mile		57,6	131,4	57,6	115,3
Fuel Costs / Mile		4,2	141,6	4,2	141,6
Profit / Passenger Mile		0,2	103,1	0,2	126,5
Revenue / Passenger		289,6	109,7	289,6	104,6

0% 100% 200%

top 10 routes (last 30 days)

#	from	to	passengers in%	profit in %
1	Los Angeles	Oakland	12,6	10,5
2	Los Angeles	Vegas	9,7	10,2
3	Oakland	Dallas	8,2	8,7
4	Dallas	Houston	6,3	7,5
5	Oakland	Seattle	6,3	6,9
6	Houston	Orlando	3,9	4,2
7	Chicago	Dallas	2,6	3,2
8	Chicago	Orlando	2,0	2,1
9	Los Angeles	Orlando	2,1	1,8
10	Oakland	Orlando	1,9	1,7

worse 10 routes (last 6 months)

#	from	to	cancelled in%	delayed in%
1	Detroit	Orlando	5,1	31,4
2	Chicago	Dallas	4,6	26,3
3	Minneapolis	Denver	4,2	29,7
4	Houston	Orlando	4,1	21,7
5	Chicago	Orlando	3,9	25,6
6	Memphis	Detroit	3,2	15,8
7	Salt Lake City	Boston	2,8	19,7
8	Oakland	Orlando	1,9	14,9
9	Dallas	Houston	1,1	16,7
10	Oakland	Seattle	0,9	14,3

cancel./delays by reason (last 30 days)

#	reason	cancelled	delayed
1	Weather	6	76
2	Missing or late flight crew	2	17
3	Mechanical failure	1	15
4	Missing or late ground crew	1	4
5	Inefficient gate handling	0	2
6	Other	2	3

International Credentialing Associates

Data Visualization Competition

ATP Innovations in Testing 2014
Scottsdale, AZ

Process

1. Announcement
2. Purpose
3. Logistics
4. Requirements
5. Evaluation

ATP Innovations in Testing 2014
Scottsdale, AZ

Competition Announcement

April 15, 2013

LinkedIn

AMERICAN EVALUATION ASSOCIATION

International Credentialing Associates

Purpose

Pretest(P_/Score(S)	N	P+	Rbis	Response	ABC*	ABD	ACD	BCD	Skipped	Flag: P+	Flag: Rbis	Flag: Key	---	---
P	215	0.05	-0.06	Upper Third	4 (5.4%)	51 (68.9%)	6(8.1%)	13 (17.6%)	0 (0.0%)	TOO HARD	Rbis Flag	Key Flag		
				Middle Third	4 (6.3%)	34 (54%)	10 (15.9%)	15 (23.8%)	0 (0.0%)					
				Lower Third	3 (3.8%)	42 (53.8%)	24 (30.8%)	9 (11.5%)	0 (0.0%)					
				Total	**11 (5.2%)**	**127 (59.6%)**	**40 (18.8%)**	**37 (16.4%)**	**0 (0.0%)**					

Revise the Current Item Performance Report to

- Improve the look
- Increase communication ease

Logistics

Rules of Entry

Winner Awarded

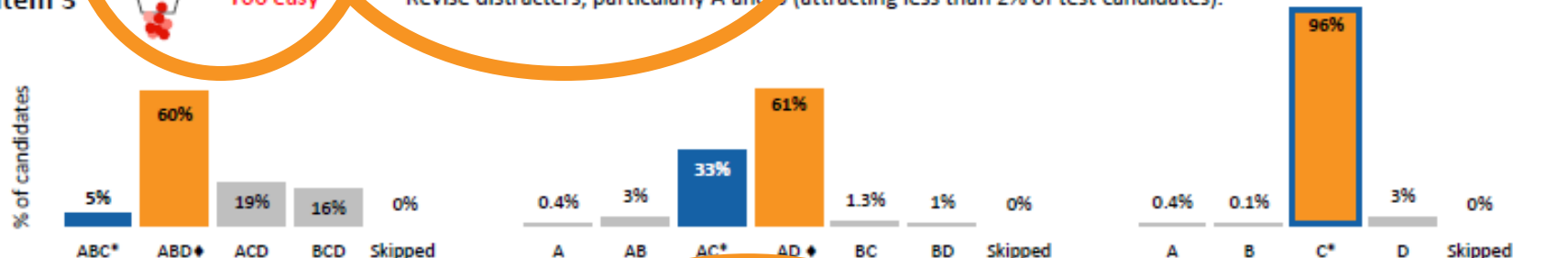
Reporting Requirements

ATP Innovations in Testing 2014
Scottsdale, AZ

Confirmation of Test Keys

This report presents the findings, recommendations and data for three test items. The purpose of the analysis is to identify which test keys to include as distractors or if the test key is in need of revision to attract candidates to distractors as they select their answer. Included is data of how Upper, Middle and Lower third performers answered and the item's analysis score.

Findings		Recommendations
Item 1	Too hard	Confirm key of ABD and review option A C to ensure that it is incorrect. Statistics suggest that ABD is key.
Item 2	Acceptable	Confirm key of AC and review option AD o ensure that it is incorrect. Statistics suggest that AD is key.
Item 3	Too easy	Revise distracters, particularly A and (attracting less than 2% of test candidates).



Item 1

Performers	ABC*	ABD♦	ACD	BCD	TOTAL	Skips
Upper	5%	69%	8%	18%	34%	0%
	4	51	6	13	74	0
Middle	6%	54%	16%	24%	30%	0%
	4	34	10	15	63	0
Lower	4%	54%	31%	12%	36%	0%
	3	42	24	9	78	0
					100%	0%
	P+ = 0.05		Rbis = -0.0		215	0

Item 2

	A	AB	AC*	AD♦	BC	BD	TOTAL	Skips
Upper	0%	0%	39%	61%	0%	0%	42%	0%
	0	0	37	58	0	0	95	0
Middle	0%	8%	26%	65%	0%	0%	2%	0%
	0	6	19	47	0	0	72	0
Lower	2%	2%	33%	56%	5%	3%	28%	0%
	3	1	20	34	3	2	63	0
							100%	0%
	P+ =		Rbis =				228	0

Item 3

	A	B	C*	D	TOTAL	Skips
Upper	0%	0%	99.6%	0.4%	32%	0%
	0	0	261	1	262	0
Middle	0%	0%	98.1%	2%	33%	0%
	0	0	263	5	268	0
Lower	0.2%	0.3%	92.5%	6%	34%	0%
	3	1	258	17	279	0
					100%	0%
	P+ = 0.96		Rbis = 0.26		809	0

*Correct option

♦ Option that complet the form with the correct answer (key).

the item is too easy. If < 0.30, the item is too hard.

Rbis value: Correlation between performance on an item and total performance. A value less than .20 is considered to not be discriminating well. In other words, the higher performers tend to answer this item incorrectly and the lower performers tend to answer incorrectly.

Prepar d for CUSTOMER NAME, April 2 13

International Credentialing Associates

Evaluation Criteria	0 Points	1 Point	2 Points	3 Points	4 Points
Uses basic spreadsheet software (Excel, Open Office, etc.) for data visualizations	Does not use basic spreadsheet software	x	x	Uses basic spreadsheet software	
All required components are present	One or more is missing	All are present	X	X	X
All components fit on one page	Does not fit on one page	Fits on one page	X	X	X
Text is legible	No text is legible when printed on 8.5 x 11 paper	Some, but not all, text is legible	All text is legible	X	X
Data visualizations are easily interpreted	Even those who know the data well cannot understand what several graphs are saying	Even those who know the data well cannot understand what one or two graphs are saying but the rest are interpretable	With insider knowledge, someone could interpret the data visualizations	With a little effort and a little knowledge of ICA, someone could interpret the data visualizations	Someone unfamiliar with the data could interpret the data visualizations
Transfer of graphs from spreadsheet to report is easy	Impossible to implement without original designer	Can be implemented without original designer but requires substantial training	Can be implemented independently but requires some experience	Can be implemented independently but requires a substantial amount of time but minimal experience is required	Can be implemented independently but requires very little time no experience is required
Report uses best practices in data visualization design					
Layout and use of space	Report is very cluttered and feels disorganized	Report is somewhat cluttered and feels busy	Report is uncluttered and feels clean	X	X
Use of color and contrast	Colors clash or do not contrast well, do not use ICA color scheme	Colors work okay but does not use ICA color scheme	Colors contrast well, use ICA color scheme	X	X

Results

International Credentialing Associates

Acme Corporation

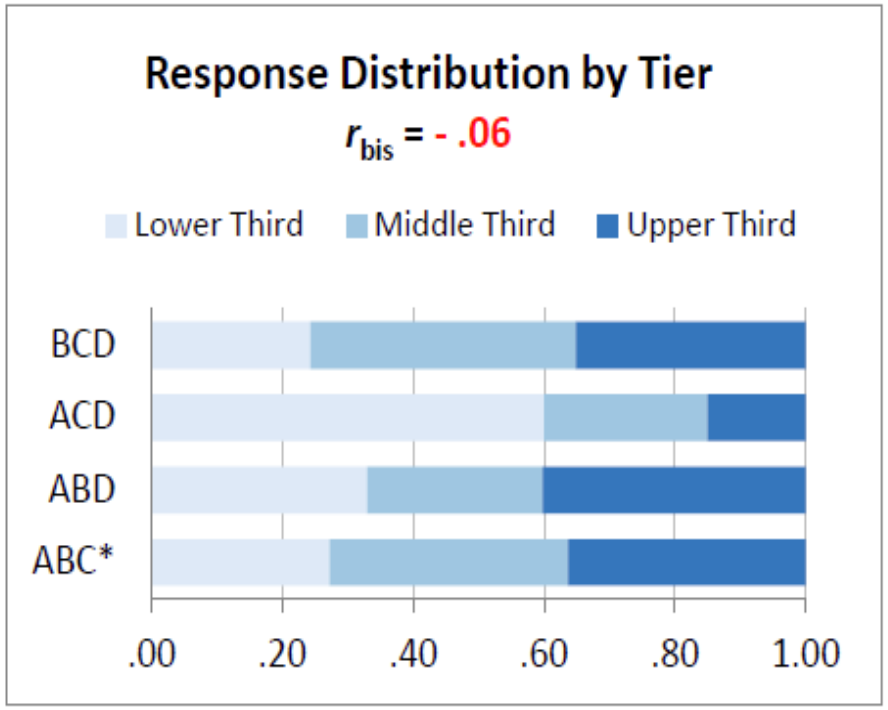
Item Name	Rbis	N	P+		Response	Upper Third	Middle Third	Lower Third	Total	
XXXX Multiple Resp.	-0.06	215	5%	too hard	ABC	4	4	3	11	5%
					ABD	51	34	42	127	59%
					ACD	6	10	24	40	19%
Confirm key of ABD and review option ABC to ensure that it is incorrect. Statistics are suggesting that ABD is key.					BCD	13	15	9	37	17%
					Skipped	0	0	0	0	0%
XXXXX Multiple Resp.	-0.09	228	33%		A	0	0	1	1	0%
					AB	0	6	1	7	3%
Confirm key of AC and review option AD to ensure that it is incorrect. Statistics are suggesting that AD is key.					AC	37	19	20	76	33%
					AD	58	47	34	139	61%
					BC	0	0	3	3	1%
					BD	0	0	2	2	1%
					Skipped	0	0	0	0	0%
XXXXX	0.26	809	96%	too easy	A	0	0	3	3	0%
					B	0	0	1	1	0%
Revise distracters, particularly A and B (attracting less than 2% of test candidates)					C	261	263	258	782	97%
					D	1	5	17	23	3%
					Skipped	0	0	0	0	0%

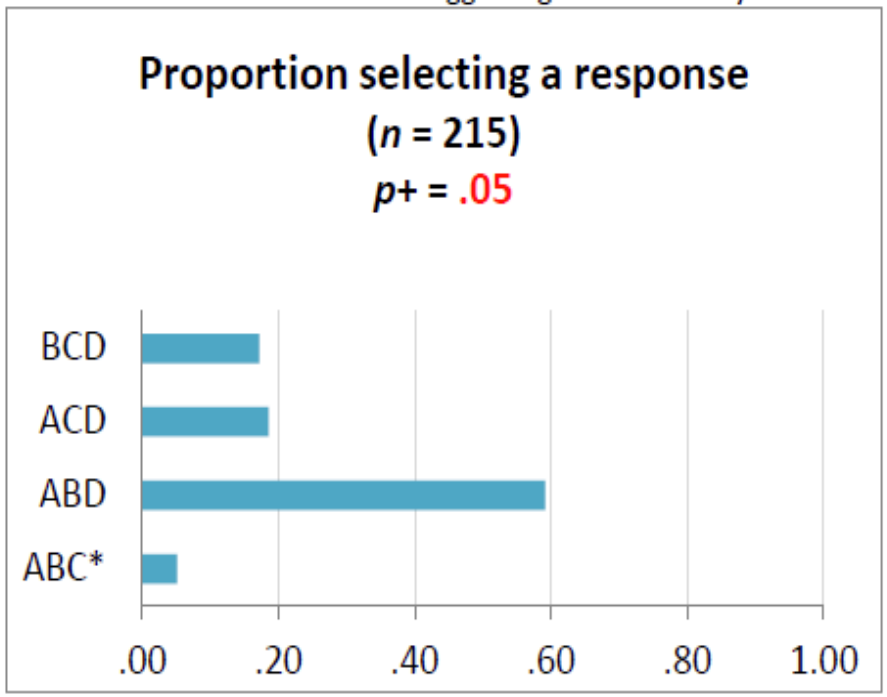
International Credentialing Associates

Customer Name

Item X1

Recommended Actions: Confirm key of ABD and review option ABC to ensure that it is incorrect. Statistics are suggesting that ABD is key.





TEST ITEMS PERFORMANCE

Organisation

THE EXAMPLE FOUNDATION

Item

XXX Multiresponse

Key

There is an option that completes the form with the correct answer.

N

Number of candidates attended the item — 215

P+

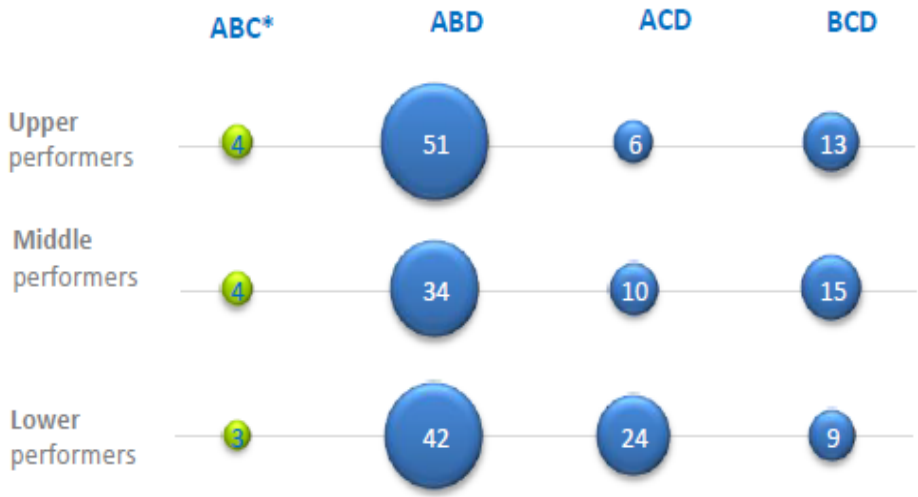
Number of candidates answering correctly — 5%

If <30%, too hard; if >92%, too easy

Rbis

Correlation between performance in this item and total performance - 0,06

If <0,20, correlation poor, not discriminating well.



Skipped = 0%

RECOMMENDATIONS

Confirm key of ABD and review option ABC to ensure that it is incorrect. Statistics are suggesting that ABD is key.

ATP Innovations in Testing 2014 Scottsdale, AZ

International Credentialing Associates

TEST ITEMS PERFORMANCE

Organisation

THE EXAMPLE FOUNDATION

Item

Key

XXX Multiresponse

There is an option that completes the form with the correct answer.

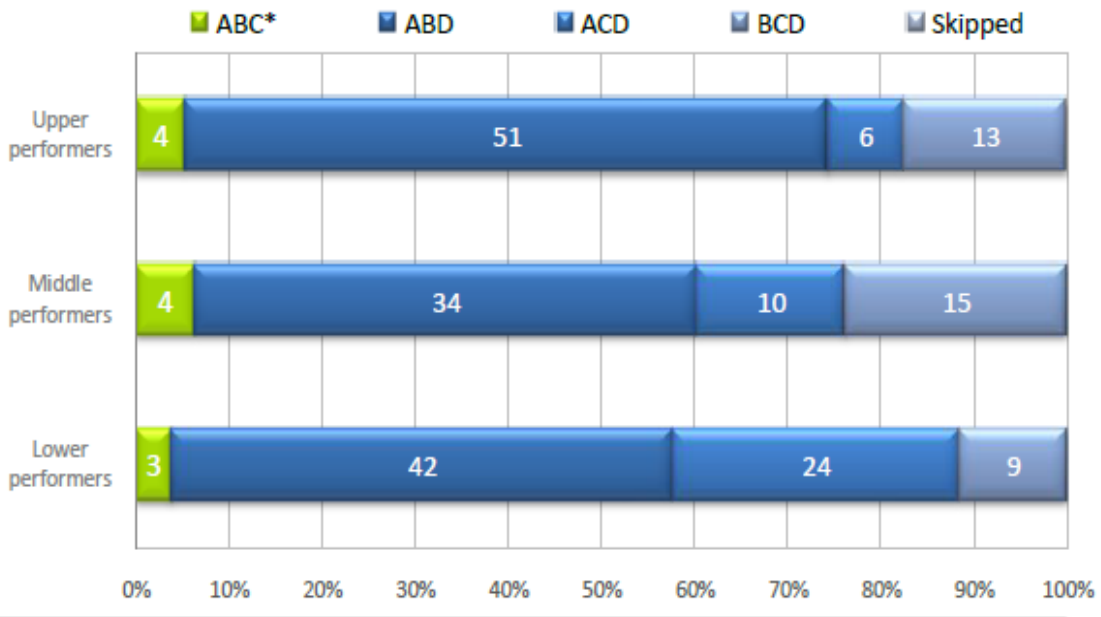
Number of candidates attended the item: 215 (N)

Number of candidates answering correctly: 5% (P+)

If <30%, too hard; if >92%, too easy

Correlation between performance in this item and total performance: - 0,06 (Rbis)

If <0,20, correlation poor, not discriminating well.



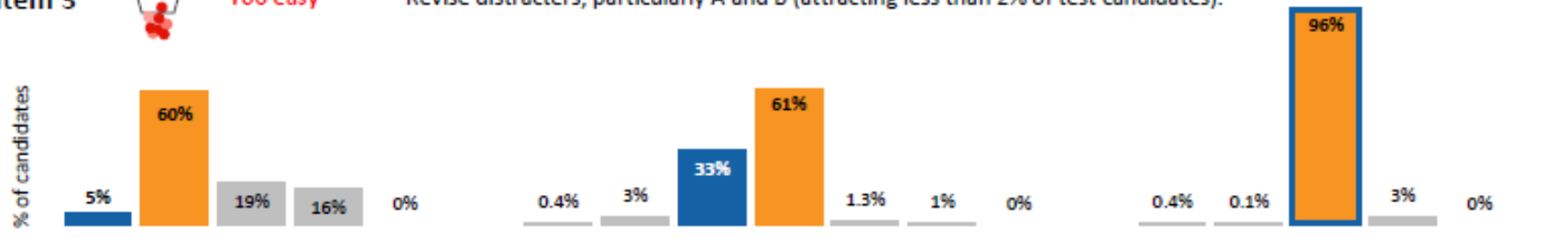
RECOMMENDATIONS

Confirm key of ABD and review option ABC to ensure that it is incorrect. Statistics are suggesting that ABD is key.

Confirmation of Test Keys

This report presents the findings, recommendations and data for three test items. The purpose of the analysis is to identify which test keys to include as distractors or if the test key is in need of revision to attract candidates to distractors as they select their answer. Included is data of how Upper, Middle and Lower third performers answered and the item's analysis score.

Findings		Recommendations
Item 1	Too hard	Confirm key of ABD and review option ABC to ensure that it is incorrect. Statistics suggest that ABD is key.
Item 2	Acceptable	Confirm key of AC and review option AD to ensure that it is incorrect. Statistics suggest that AD is key.
Item 3	Too easy	Revise distracters, particularly A and B (attracting less than 2% of test candidates).



	Item 1						Item 2								Item 3					
Performers	ABC*	ABD♦	ACD	BCD	TOTAL	Skips	A	AB	AC*	AD♦	BC	BD	TOTAL	Skips	A	B	C*	D	TOTAL	Skips
Upper	5%	69%	8%	18%	34%	0%	0%	0%	39%	61%	0%	0%	42%	0%	0%	0%	99.6%	0.4%	32%	0%
	4	51	6	13	74	0	0	0	37	58	0	0	95	0	0	0	261	1	262	0
Middle	6%	54%	16%	24%	30%	0%	0%	8%	26%	65%	0%	0%	32%	0%	0%	0%	98.1%	2%	33%	0%
	4	34	10	15	63	0	0	6	19	47	0	0	72	0	0	0	263	5	268	0
Lower	4%	54%	31%	12%	36%	0%	2%	2%	33%	56%	5%	3%	28%	0%	0.2%	0.3%	92.5%	6%	34%	0%
	3	42	24	9	78	0	3	1	20	34	3	2	63	0	3	1	258	17	279	0
					100%	0%							100%	0%					100%	0%
	P+ = 0.05		Rbis = -0.06		215	0		P+ = 0.33		Rbis = -0.09			228	0	P+ = 0.96		Rbis = 0.26		809	0

*Correct option

♦ Option that completes the form with the correct answer (key).

P+ value: > 0.92, the item is too easy. If < 0.30, the item is too hard.

Rbis value: Correlation between performance on an item and total performance. A value less than .20 is considered to not be discriminating well. In other words, the higher performers tend to answer this item incorrectly and the lower performers tend to answer incorrectly.

Prepared for CUSTOMER NAME, April 2013

International Credentialing Associates

Guess Who Won?

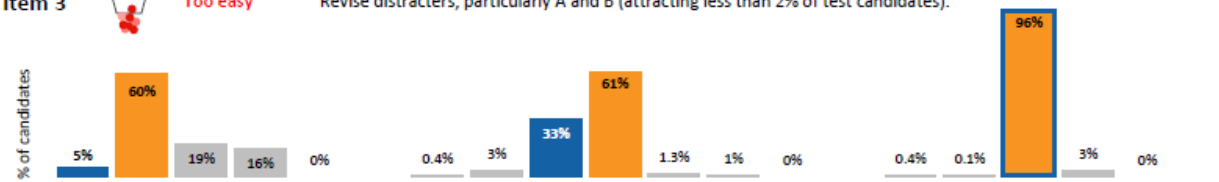
Winner is…

Jill Lipski Cain
Senior Research Analyst
The Improve Group

Confirmation of Test Keys

This report presents the findings, recommendations and data for three test items. The purpose of the analysis is to identify which test keys to include as distractors or if the test key is in need of revision to attract candidates to distractors as they select their answer. Included is data of how Upper, Middle and Lower third performers answered and the item's analysis score.

Findings		Recommendations
Item 1	Too hard	Confirm key of ABD and review option ABC to ensure that it is incorrect. Statistics suggest that ABD is key.
Item 2	Acceptable	Confirm key of AC and review option AD to ensure that it is incorrect. Statistics suggest that AD is key.
Item 3	Too easy	Revise distracters, particularly A and B (attracting less than 2% of test candidates).



	Item 1						Item 2								Item 3					
Performers	ABC*	ABD♦	ACD	BCD	TOTAL	Skips	A	AB	AC*	AD♦	BC	BD	TOTAL	Skips	A	B	C*	D	TOTAL	Skips
Upper	5%	69%	8%	18%	34%	0%	0%	0%	39%	61%	0%	0%	42%	0%	0%	0%	99.6%	0.4%	32%	0%
	4	51	6	13	74	0	0	0	37	58	0	0	95	0	0	0	261	1	262	0
Middle	6%	54%	16%	24%	30%	0%	0%	8%	26%	65%	0%	0%	32%	0%	0%	0%	98.1%	2%	33%	0%
	4	34	10	15	63	0	0	6	19	47	0	0	72	0	0	0	263	5	268	0
Lower	4%	54%	31%	12%	36%	0%	2%	2%	33%	56%	5%	3%	28%	0%	0.2%	0.3%	92.5%	6%	34%	0%
	3	42	24	9	78	0	3	1	20	34	3	2	63	0	3	1	258	17	279	0
					100%	0%							100%	0%					100%	0%
	P+ = 0.05		Rbis = -0.06		215	0	P+ = 0.33		Rbis = -0.09				228	0	P+ = 0.96		Rbis = 0.26		809	0

*Correct option
♦ Option that completes the form with the correct answer (key).
P+ value: > 0.92, the item is too easy. If < 0.30, the item is too hard.
Rbis value: Correlation between performance on an item and total performance. A value less than .20 is considered to not be discriminating well. In other words, the higher performers tend to answer this item incorrectly and the lower performers tend to answer incorrectly.

Prepared for CUSTOMER NAME, April 2013

International Credentialing Associates

ATP Innovations in Testing 2014 Scottsdale, AZ

Summary

- Implications
- Data Visualization Take Aways
- http://intlcred.wordpress.com/

ATP 2014
Innovations in Testing
Scottsdale, AZ

Contact

Dataviz

Stephanie

Evergreen Data
stephanie@
evergreenevaluation.com
269.425.1650

Credentialing

Alex

ICA
alex@
intlcred.com
347.907.4326

Manny

ICA
manny@
intlcred.com
240.308.8075