Module 3:
General Linear Model

MSIR 525

October 14-28, 2019

Recap of Module 2 (check list from syllabus; see pages 1-2)

- We learned about several issues in data sets (e.g., outliers, missing data, non-normal distributions) that may bring into question the robustness of empirical results
- We developed R code that will estimate descriptive statistics for a set of data
- We learned about the importance of interpreting and communicating descriptive statistics (e.g., in tandem, visually and empirically)
- Although we did not perform an ANOVA to assess if means differed across multiple groups, we discuss the technique's utility and limitations
- We learned how to perform a t-test; interpret its results; use its results to inform an evidence-based management decision
  - Importantly, we learned how to "explore further" to gain a better understanding of what the data are telling us

Agenda for Module 3

- 10/14/2019
  - Review of hackathon exercise; introduction to the general linear model (GLM); an assessment of the GLM assumptions

Agenda for Module 3

- 10/14/2019
  - Review of hackathon exercise; introduction to the general linear model (GLM); an assessment of the GLM assumptions
- 10/16/2019
  - Procedures to assess the relation between a predictor and a continuous outcome variable

Agenda for Module 3

- 10/14/2019
  - Review of hackathon exercise; introduction to the general linear model (GLM); an assessment of the GLM assumptions
- 10/16/2019
  - Procedures to assess the relation between a predictor and a continuous outcome variable
- 10/21/2019
  - Procedures to assess the relation between a predictor and a dichotomous outcome variable

Agenda for Module 3

- 10/14/2019
  - Review of hackathon exercise; introduction to the general linear model (GLM); an assessment of the GLM assumptions
- 10/16/2019
  - Procedures to assess the relation between a predictor and a continuous outcome variable
- 10/21/2019
  - Procedures to assess the relation between a predictor and a dichotomous outcome variable
- 10/23/2019
  - Module 3 recap and software tutorial

Agenda for Module 3

- 10/14/2019
  - Review of hackathon exercise; introduction to the general linear model (GLM); an assessment of the GLM assumptions
- 10/16/2019
  - Procedures to assess the relation between a predictor and a continuous outcome variable
- 10/21/2019
  - Procedures to assess the relation between a predictor and a dichotomous outcome variable
- 10/23/2019
  - Module 3 recap and software tutorial
- 10/28/2019
  - In-class exercise for credit (i.e., a hackathon)
  - Determine the strongest correlates of employee performance and turnover behavior

Agenda for Module 3

- Let’s get started! ☺

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

IMPORTANT POINT

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

IMPORTANT POINT

WE ARE NO LONGER DEALING WITH UNIVARIATE STATISTICS

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

IMPORTANT POINT

- **WE ARE NO LONGER DEALING WITH UNIVARIATE STATISTICS (MODULE 2)**

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

IMPORTANT POINT

- **WE ARE NO LONGER DEALING WITH UNIVARIATE STATISTICS (MODULE 2)**

WHAT DOES THIS MEAN?

- **MEASURES OF CENTRAL TENDENCY (E.G., MEAN) SUMMARIZE DATA PERTAINING TO JUST *ONE* VARIABLE (MODULE 2)**

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

IMPORTANT POINT

- **WE ARE NO LONGER DEALING WITH UNIVARIATE STATISTICS (MODULE 2)**
- **NOW, WE ARE DEALING WITH BIVARIATE STATISTICS (MODULE 3)**

WHAT DOES THIS MEAN?

- **MEASURES OF CENTRAL TENDENCY (E.G., MEAN) SUMMARIZE DATA PERTAINING TO JUST *ONE* VARIABLE (MODULE 2)**

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

IMPORTANT POINT

- **WE ARE NO LONGER DEALING WITH UNIVARIATE STATISTICS (MODULE 2)**
- **NOW, WE ARE DEALING WITH BIVARIATE STATISTICS (MODULE 3)**

WHAT DOES THIS MEAN?

- **MEASURES OF CENTRAL TENDENCY (E.G., MEAN) SUMMARIZE DATA PERTAINING TO JUST *ONE* VARIABLE (MODULE 2)**
- **NOW, WE ARE INTERESTED IN THE RELATION BETWEEN *TWO* VARIABLES (MODULE 3)**

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

- Effectively, you want to assess the validity of the organization's current screening tool(s)

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

- Effectively, you want to assess the validity of the organization's current screening tool(s)
  - In other words, are the screening tools useful for forecasting important outcomes that will affect organizational performance

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

How could univariate statistics be used in the aforementioned example?

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

How could univariate statistics be used in the aforementioned example?

- To summarize the central tendency of *one* variable

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

How could univariate statistics be used in the aforementioned example?

- To summarize the central tendency of *one* variable

How can bivariate statistics be used in the aforementioned example?

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

How could univariate statistics be used in the aforementioned example?

- To summarize the central tendency of *one* variable

How can bivariate statistics be used in the aforementioned example?

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

How could univariate statistics be used in the aforementioned example?

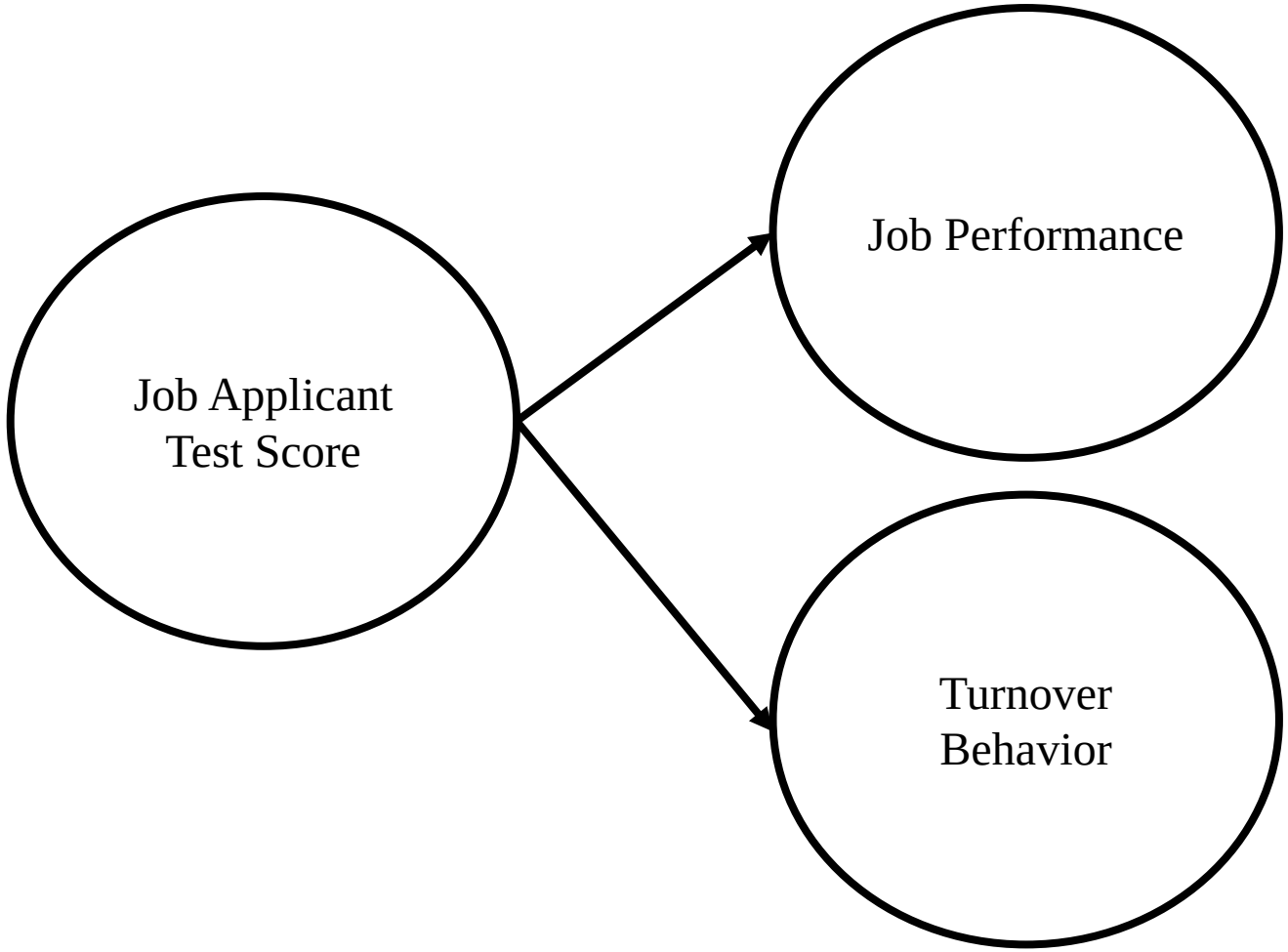
- To summarize the central tendency of *one* variable

How can bivariate statistics be used in the aforementioned example?

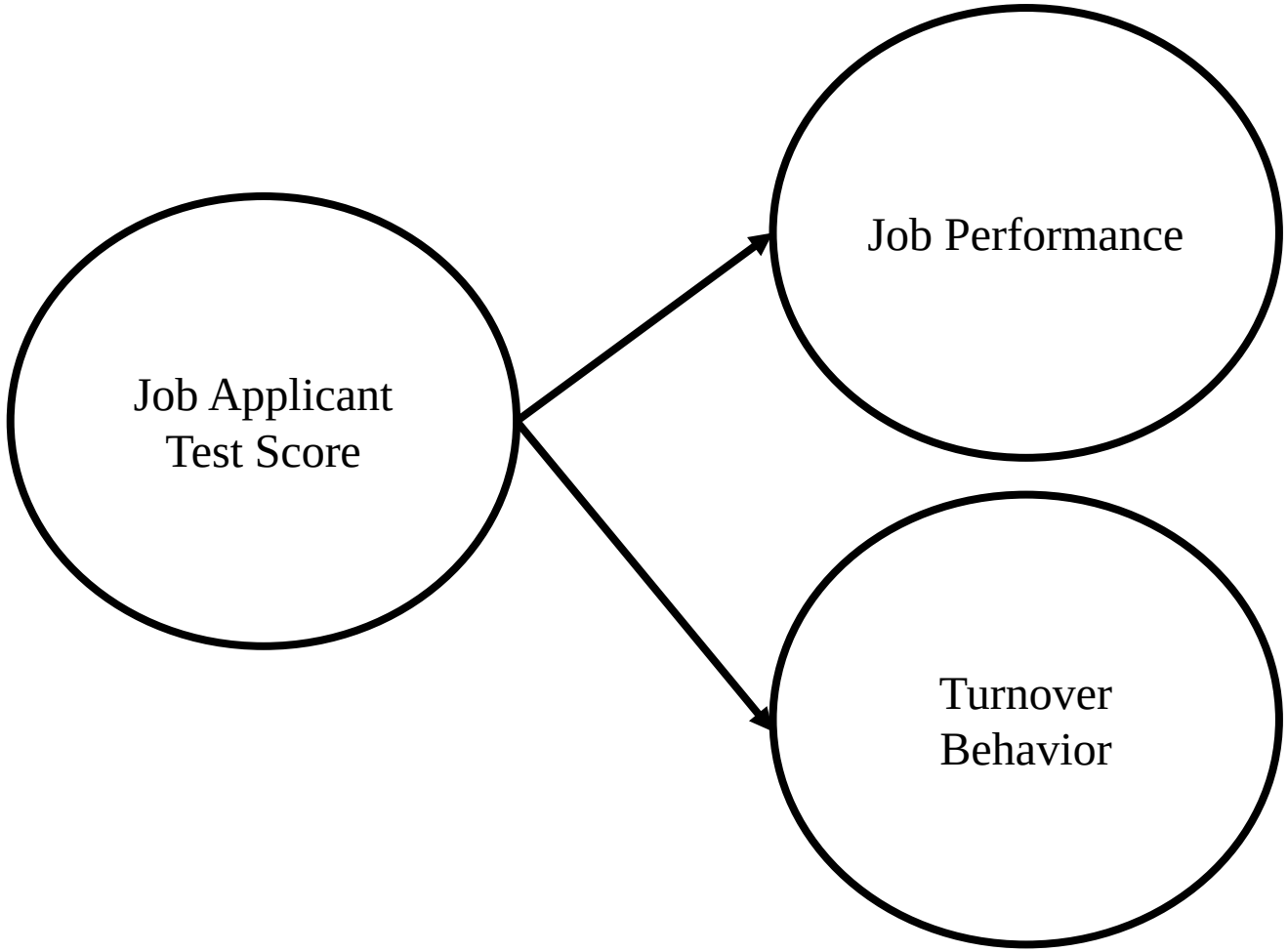
- You're right, we don't know how to do this just yet (it's the whole purpose of Module 3!

So, let's go and learn about the correlation coefficient and the simple linear regression model

Motivating Example:



Motivating Example:

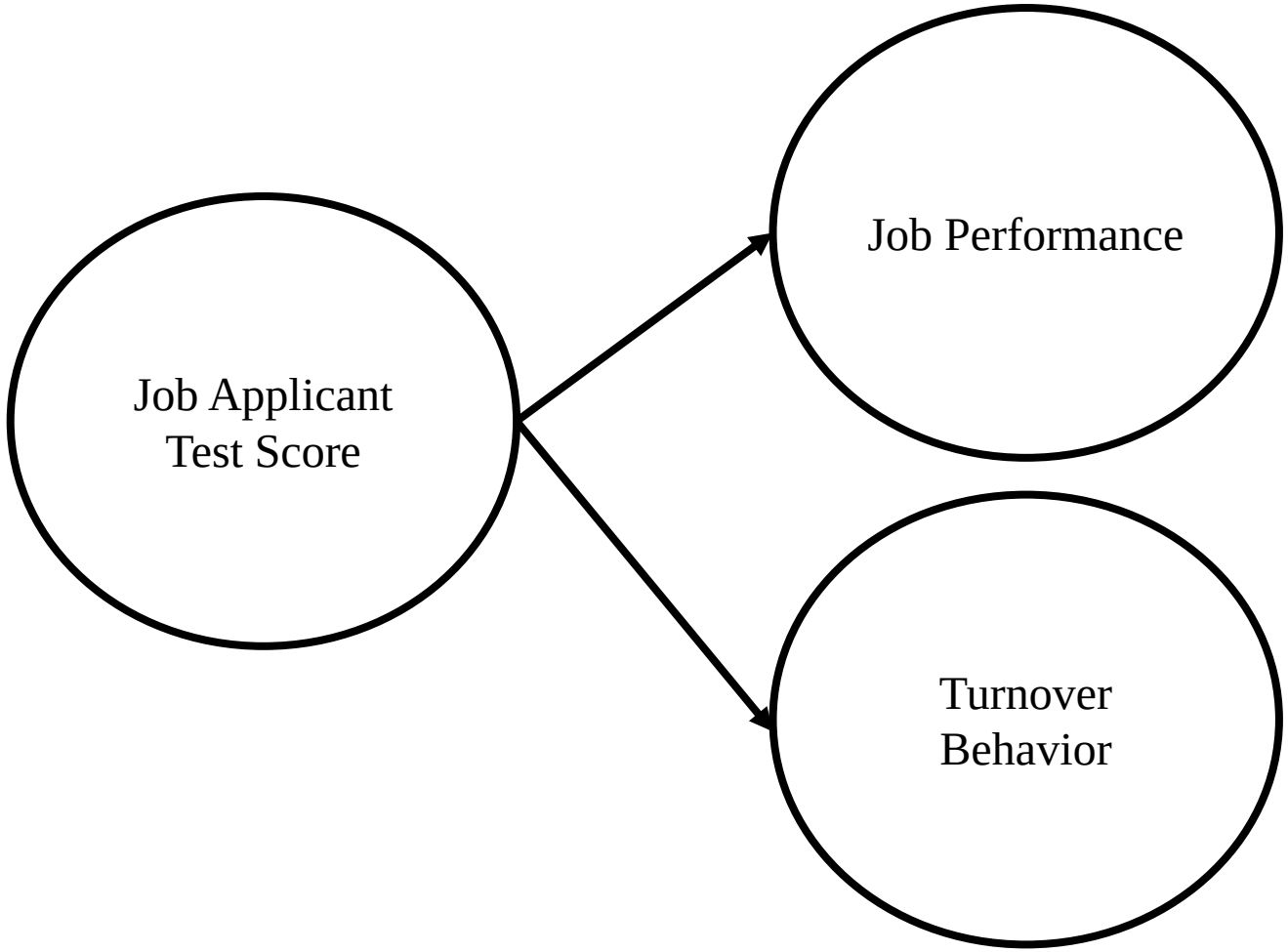


How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the association between variables in each of relation of interest

(1) Test score → Performance
(2) Test score → Turnover

Motivating Example:



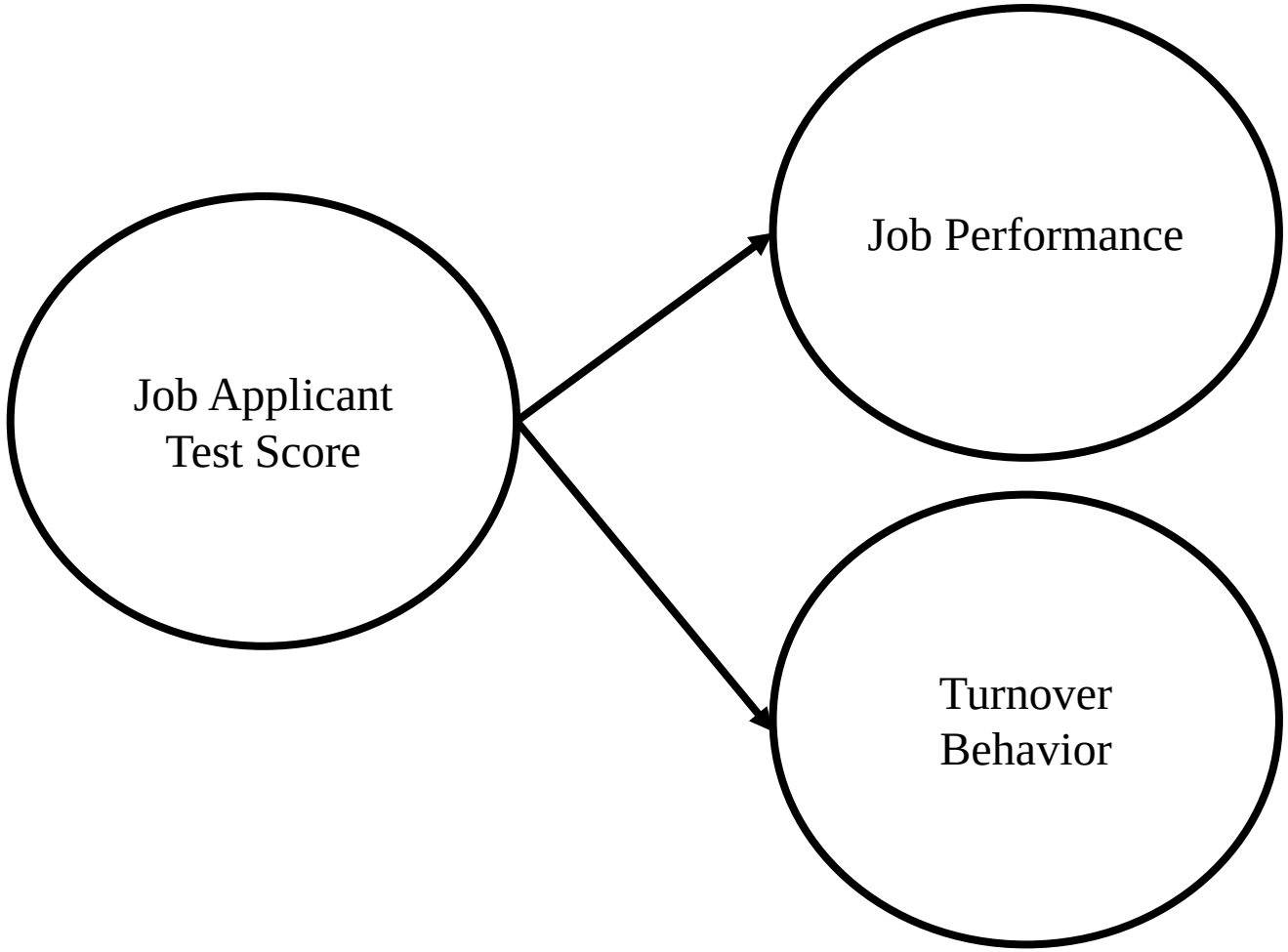
How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

(1) Test score → Performance
(2) Test score → Turnover

A subtle, but very important point, is being made here…

Motivating Example:



How can these relations be summarized?

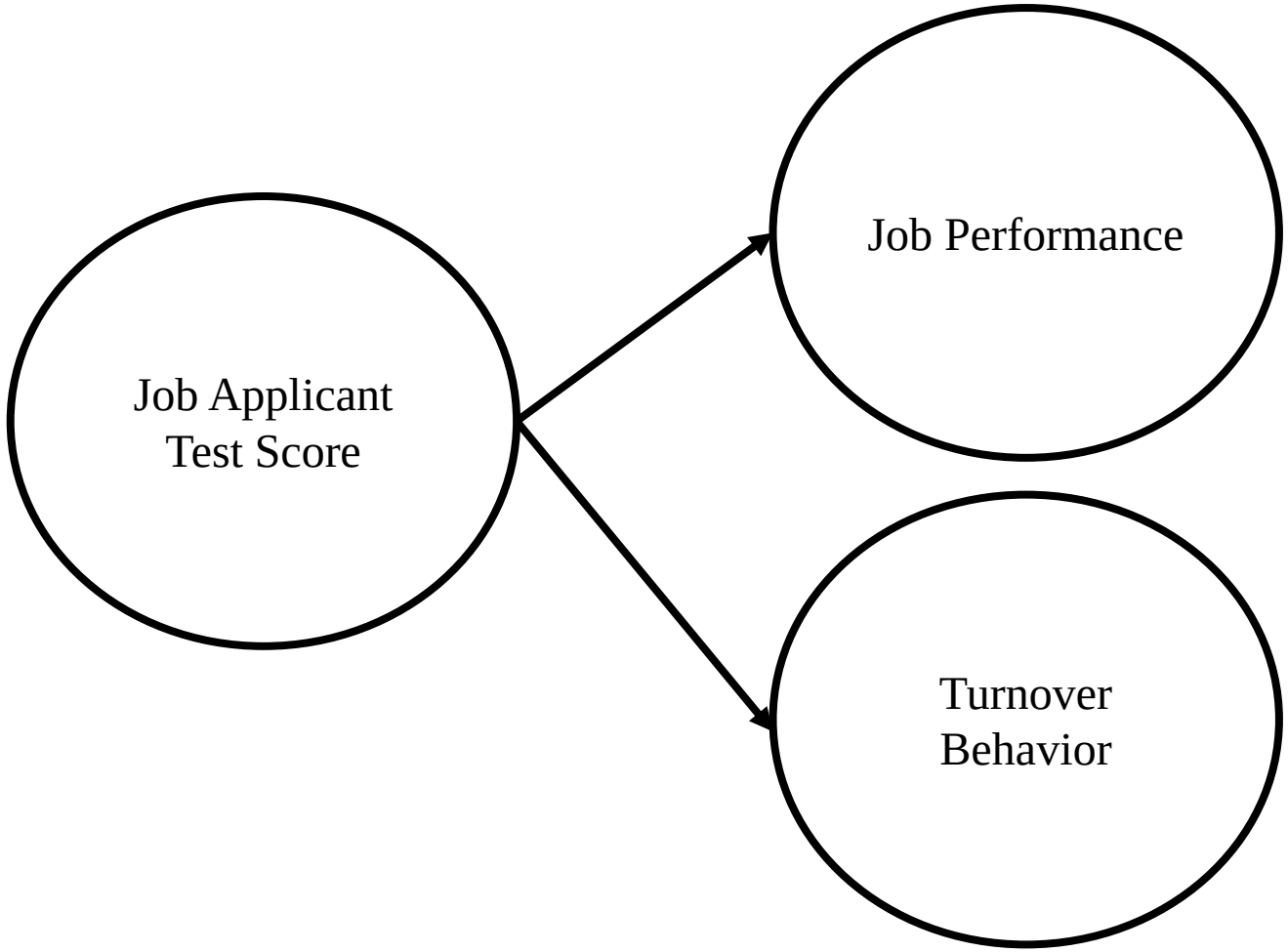
First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

(1) Test score → Performance
(2) Test score → Turnover

A subtle, but very important point, is being made here…

We are looking at the association between two things.

Motivating Example:



How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

(1) Test score → Performance
(2) Test score → Turnover

A subtle, but very important point, is being made here…

We are looking at the association between two things.

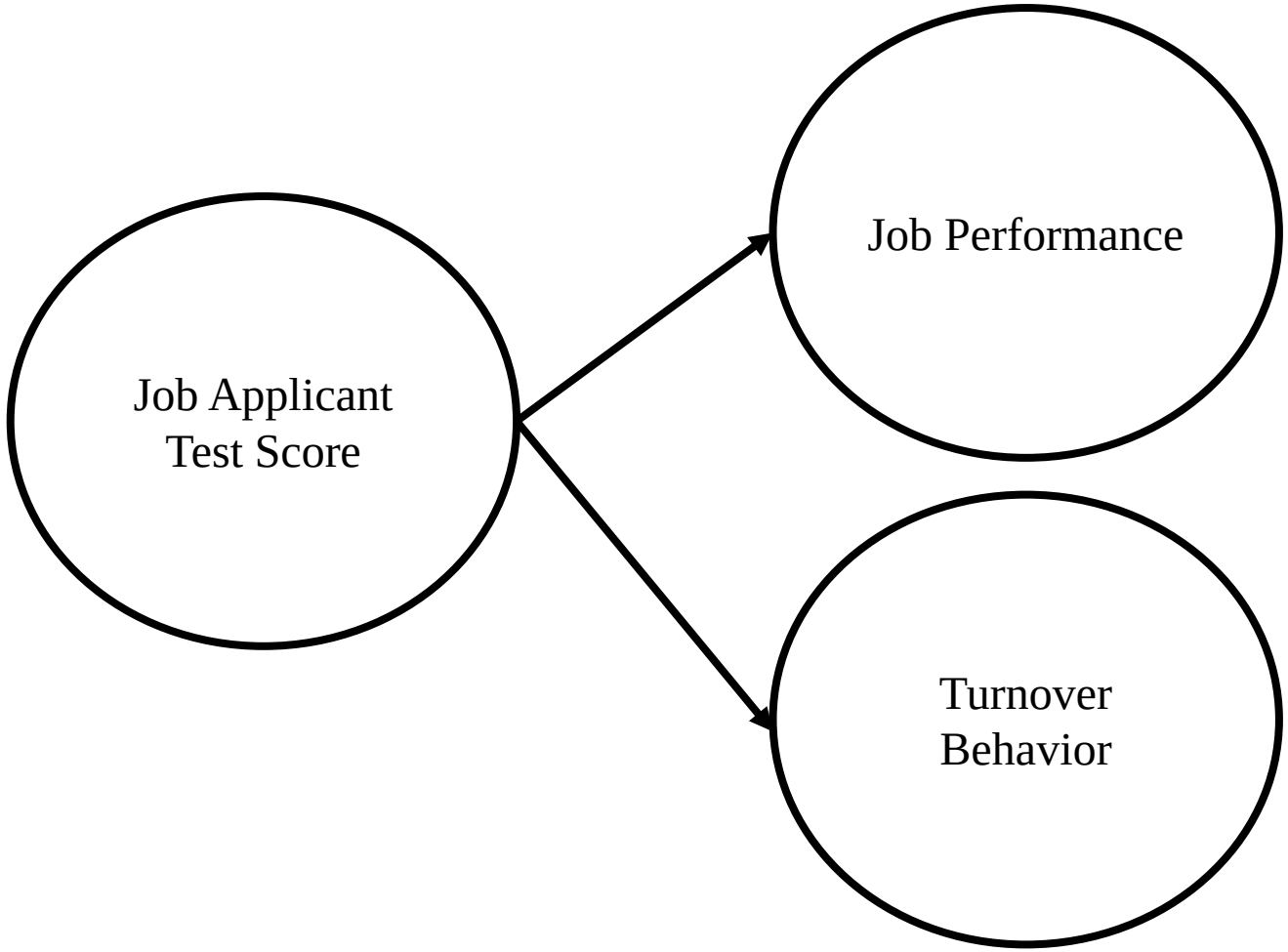
We are not predicting one them from another

Motivating Example:

- Imagine that you are an HR Analyst who is interested in knowing if there is a relationship between an individual's applicant exam score and (a) future job performance and (b) future turnover behavior.

- Effectively, you want to know if the organization's current screening tools have important **validity** outcomes.

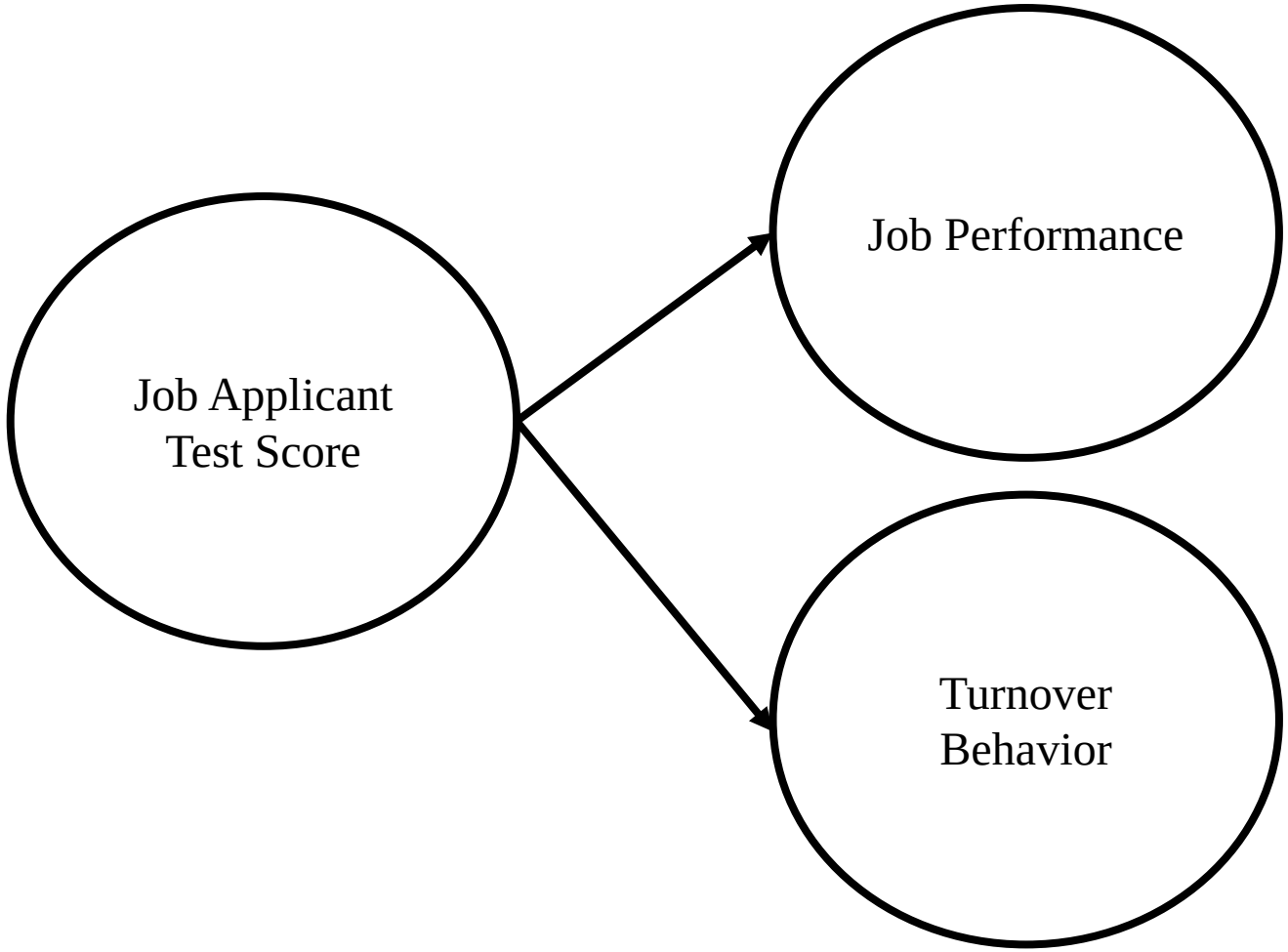
Motivating Example:



How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

Motivating Example:

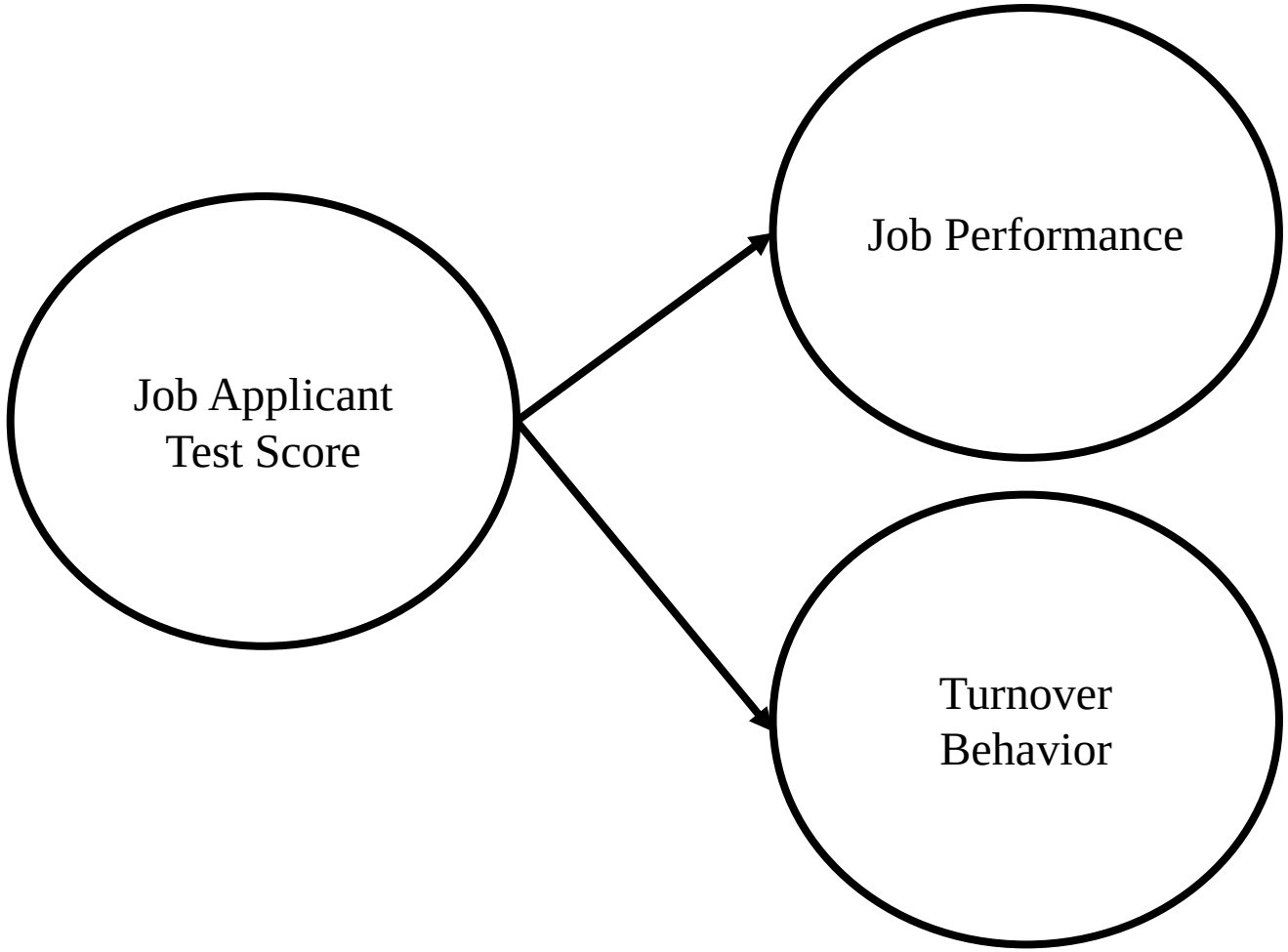


How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

The correlation coefficient has two characteristics…

Motivating Example:



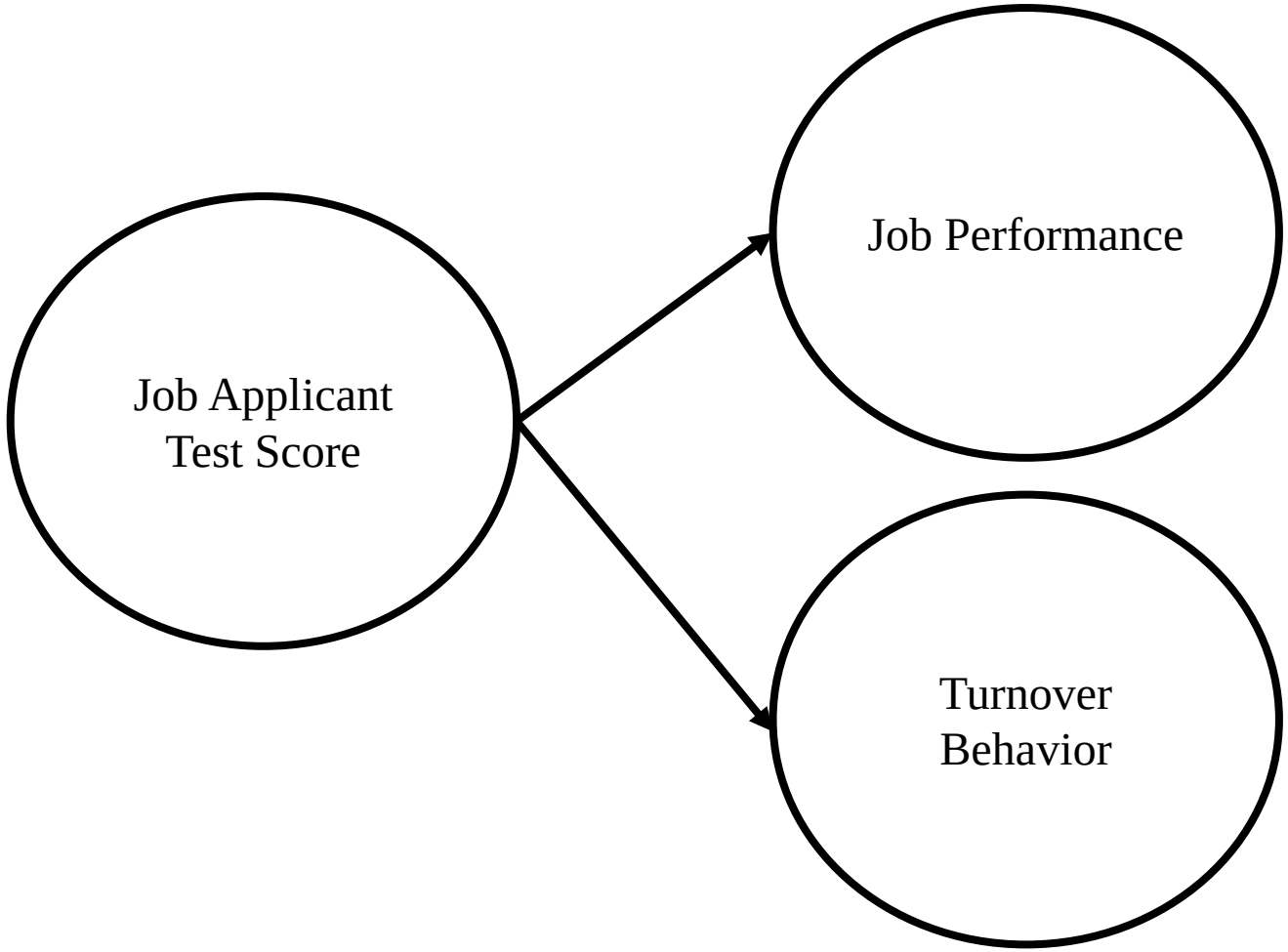
How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **<u>association</u>** between variables in each of relation of interest

The correlation coefficient has two characteristics…

(1) Direction → Tells us if the nature of the relation is positive (+) or negative (-)

Motivating Example:



How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **<u>association</u>** between variables in each of relation of interest

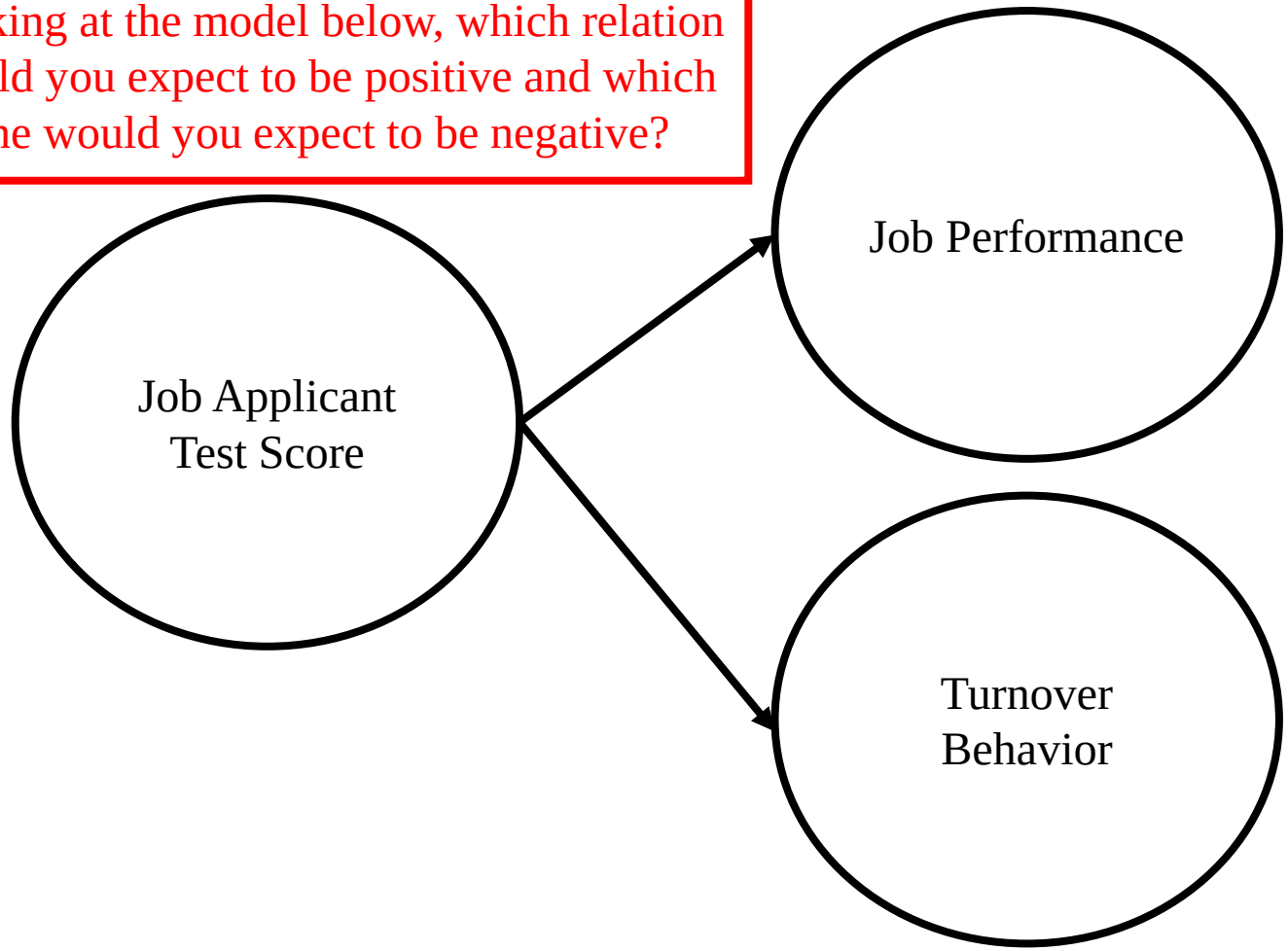
The correlation coefficient has two characteristics…

(1) Direction → Tells us if the nature of the relation is positive (+) or negative (-)

   (1) If positive…
      - As X increases, Y increases
   (2) If negative…
      - As X increases, Y decreases

Motivating Example:

Looking at the model below, which relation would you expect to be positive and which one would you expect to be negative?



How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

The correlation coefficient has two characteristics…

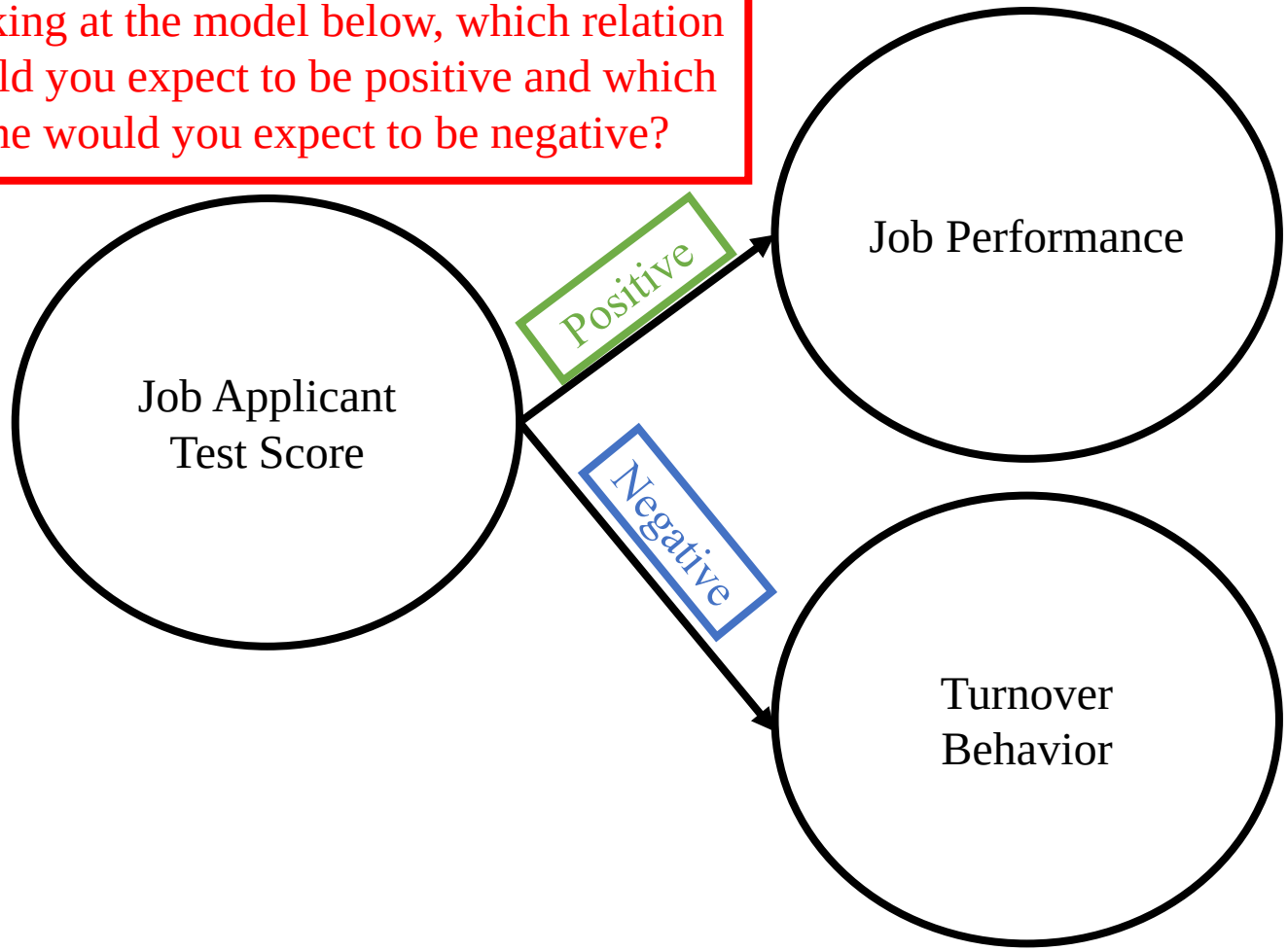
(1) Direction → Tells us if the nature of the relation is positive (+) or negative (-)

(1) If positive…
- As X increases, Y increases

(2) If negative…
- As X increases, Y decreases

Motivating Example:

Looking at the model below, which relation would you expect to be positive and which one would you expect to be negative?



How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

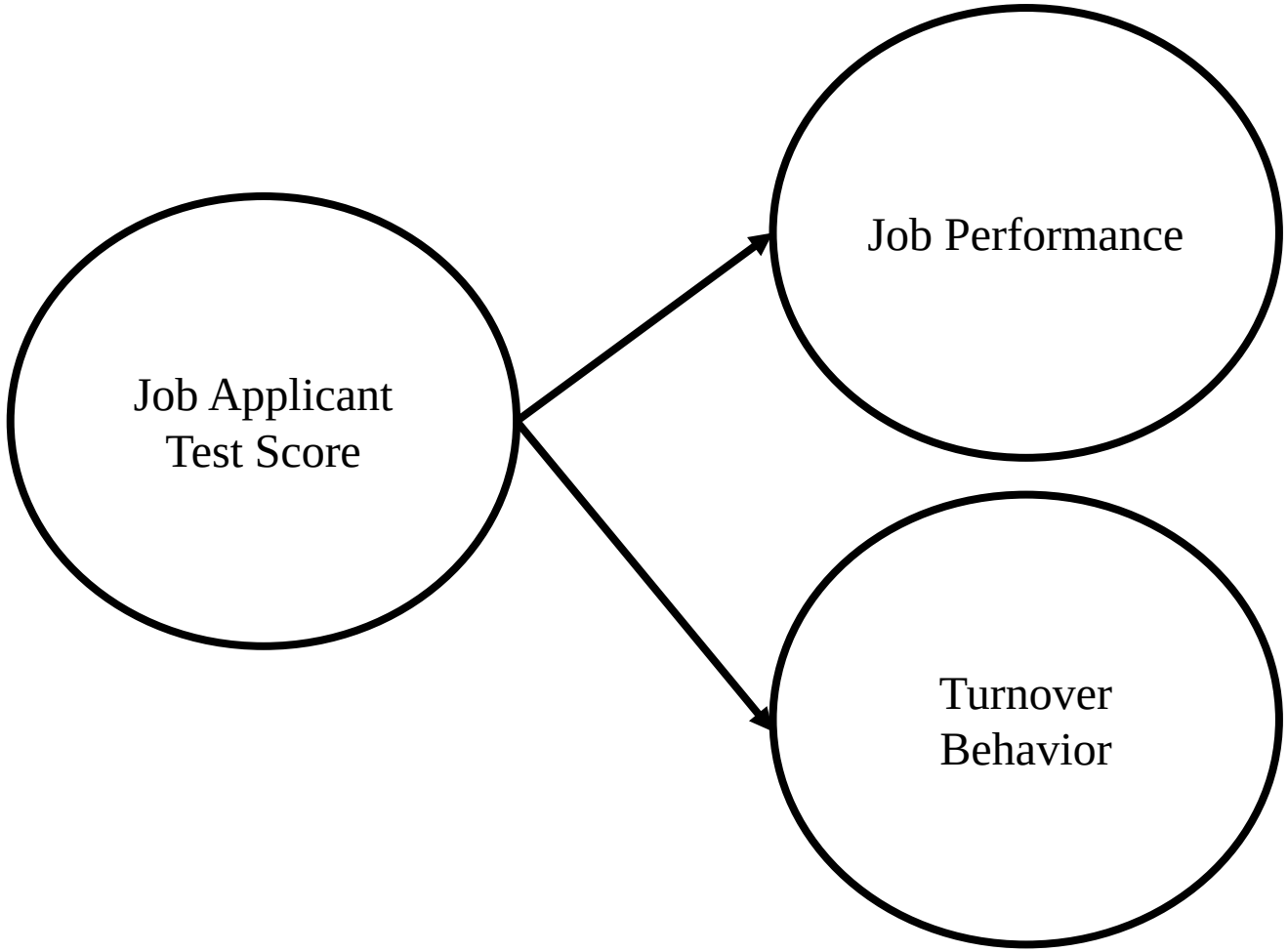
The correlation coefficient has two characteristics…

(1) Direction → Tells us if the nature of the relation is positive (+) or negative (-)

(1) If positive…
- As X increases, Y increases

(2) If negative…
- As X increases, Y decreases

Motivating Example:



How can these relations be summarized?

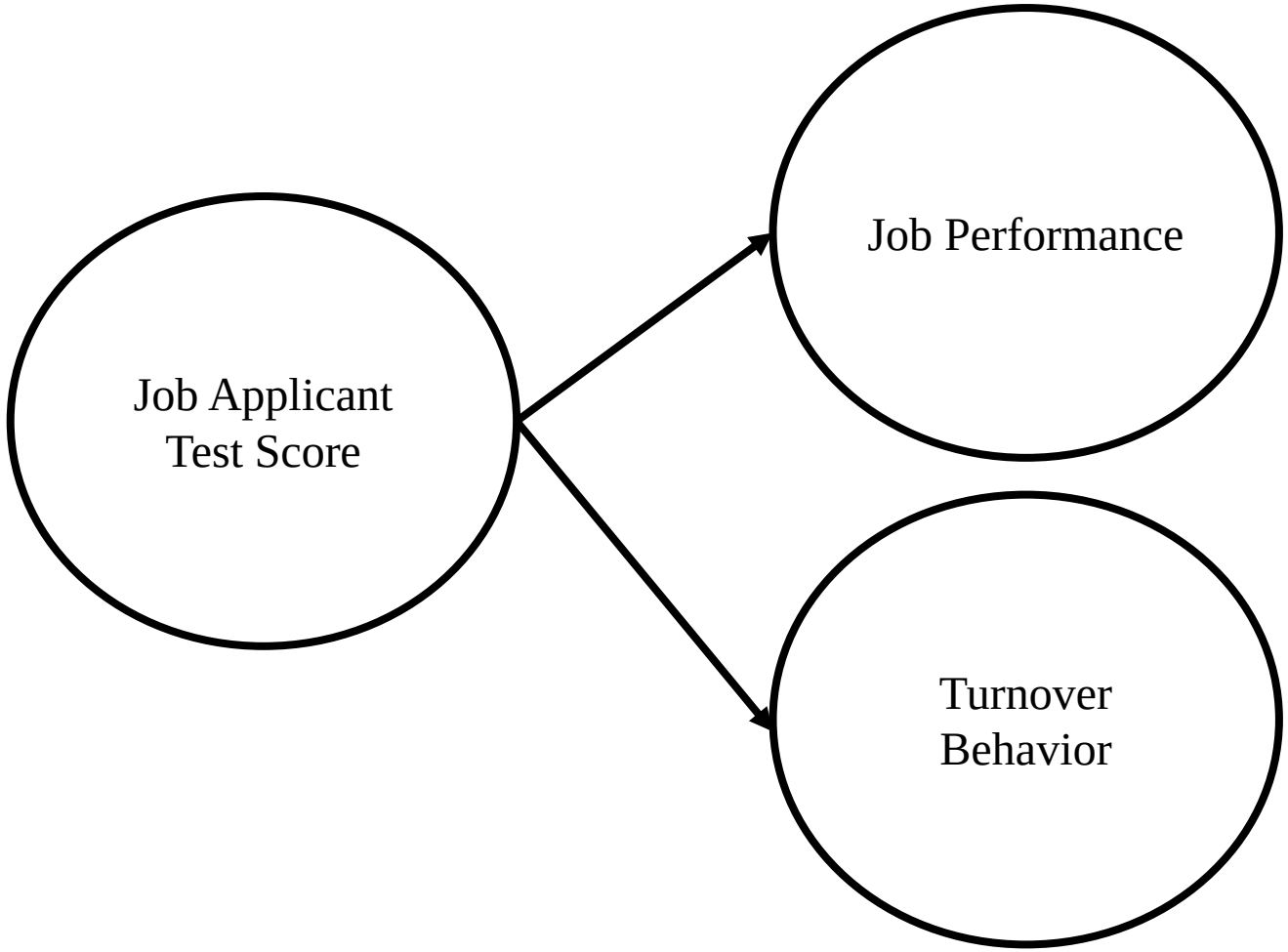
First, we can use the *correlation coefficient* to measure the **<u>association</u>** between variables in each of relation of interest

The correlation coefficient has two characteristics…

(1) Direction → Tells us if the nature of the relation is positive (+) or negative (-)

(2) Magnitude → Tells us if the relation between two things is "weak" or "strong"

Motivating Example:



How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **<u>association</u>** between variables in each of relation of interest

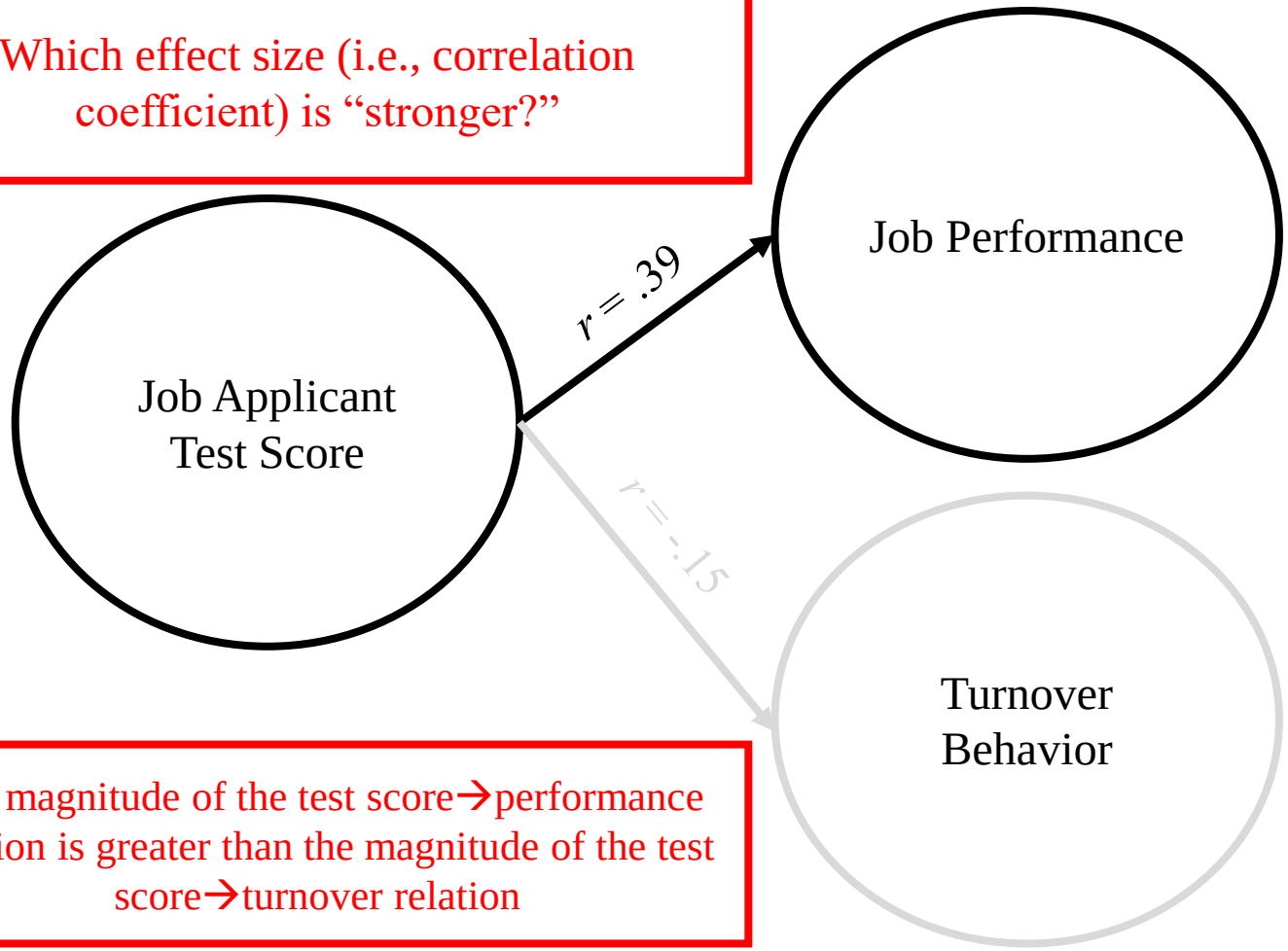
The correlation coefficient has two characteristics…

(1) Direction → Tells us if the nature of the relation is positive (+) or negative (-)

(2) Magnitude → Tells us if the relation between two things is "weak" or "strong"
   (1) Ranges on a scale from 0 → |1|

Motivating Example:

Which effect size (i.e., correlation coefficient) is "stronger?"



The magnitude of the test score→performance relation is greater than the magnitude of the test score→turnover relation

How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

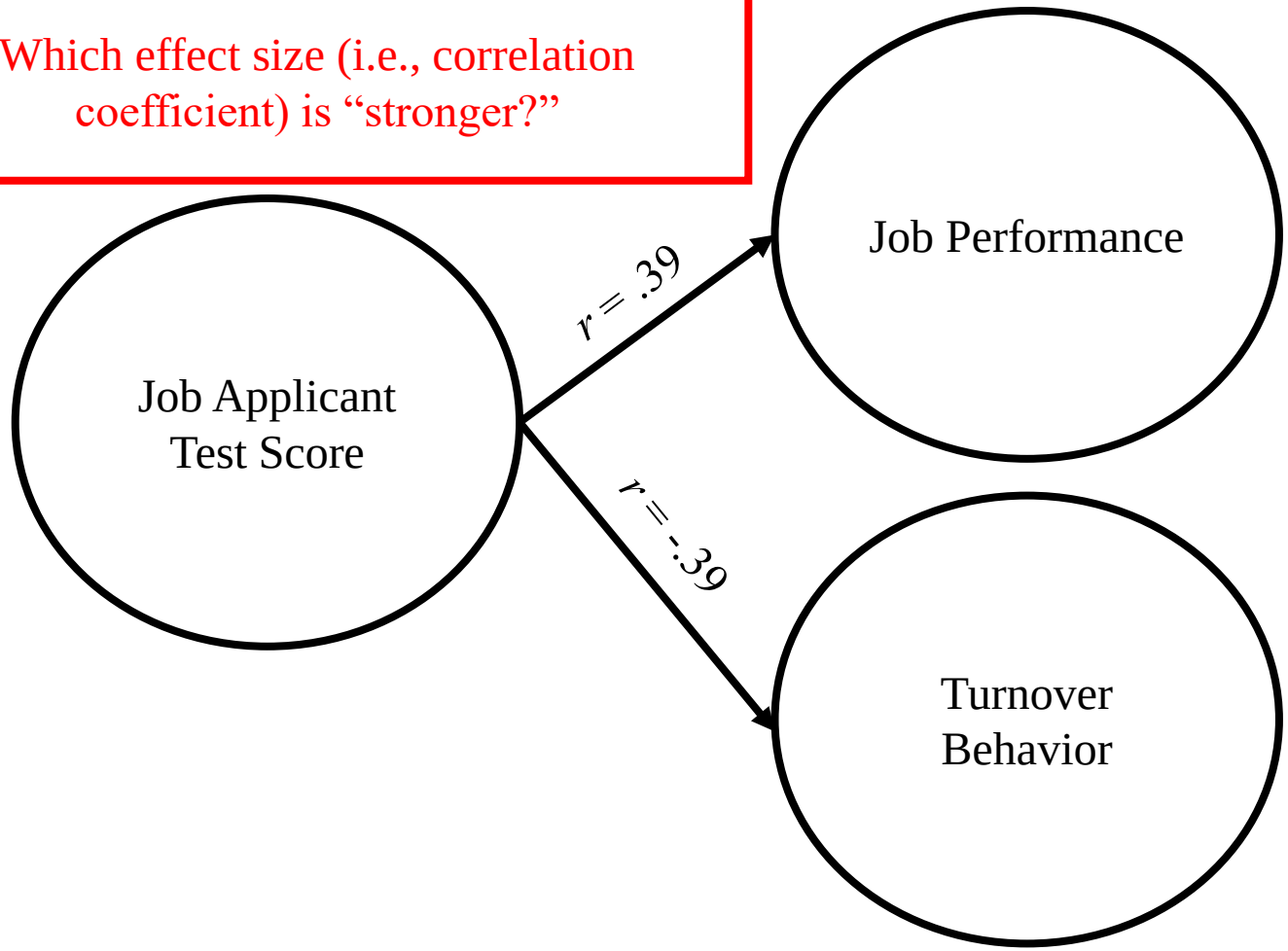
The correlation coefficient has two characteristics…

(1) Direction → Tells us if the nature of the relation is positive (+) or negative (-)

(2) Magnitude → Tells us if the relation between two things is "weak" or "strong"
   (1) Ranges on a scale from 0 → |1|

Motivating Example:

Which effect size (i.e., correlation coefficient) is "stronger?"



How can these relations be summarized?

First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

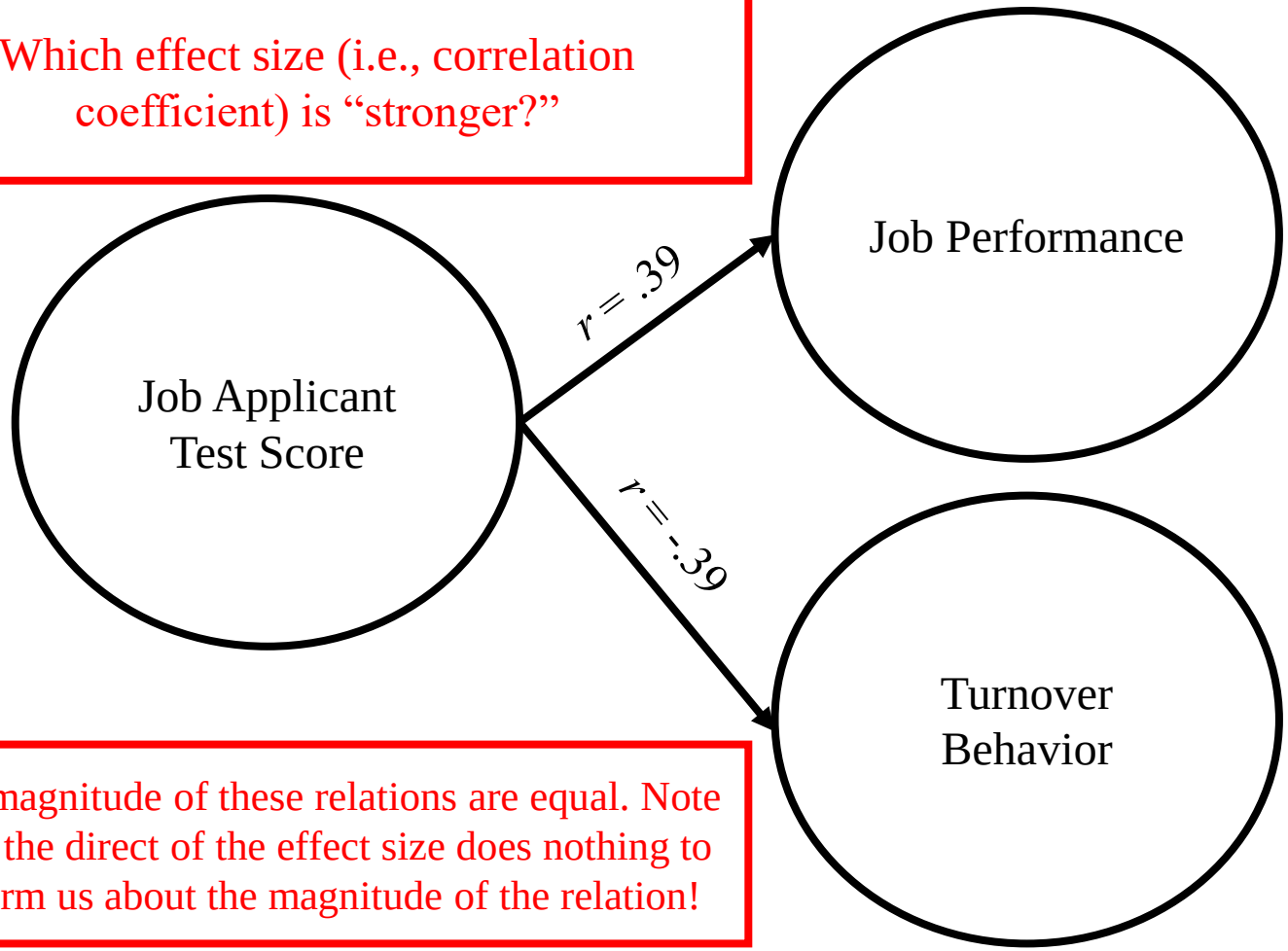
The correlation coefficient has two characteristics…

(1) Direction → Tells us if the nature of the relation is positive (+) or negative (-)

(2) Magnitude → Tells us if the relation between two things is "weak" or "strong"
   (1) Ranges on a scale from 0 → |1|

Motivating Example:

Which effect size (i.e., correlation coefficient) is "stronger?"



The magnitude of these relations are equal. Note than the direct of the effect size does nothing to inform us about the magnitude of the relation!

How can these relations be summarized?

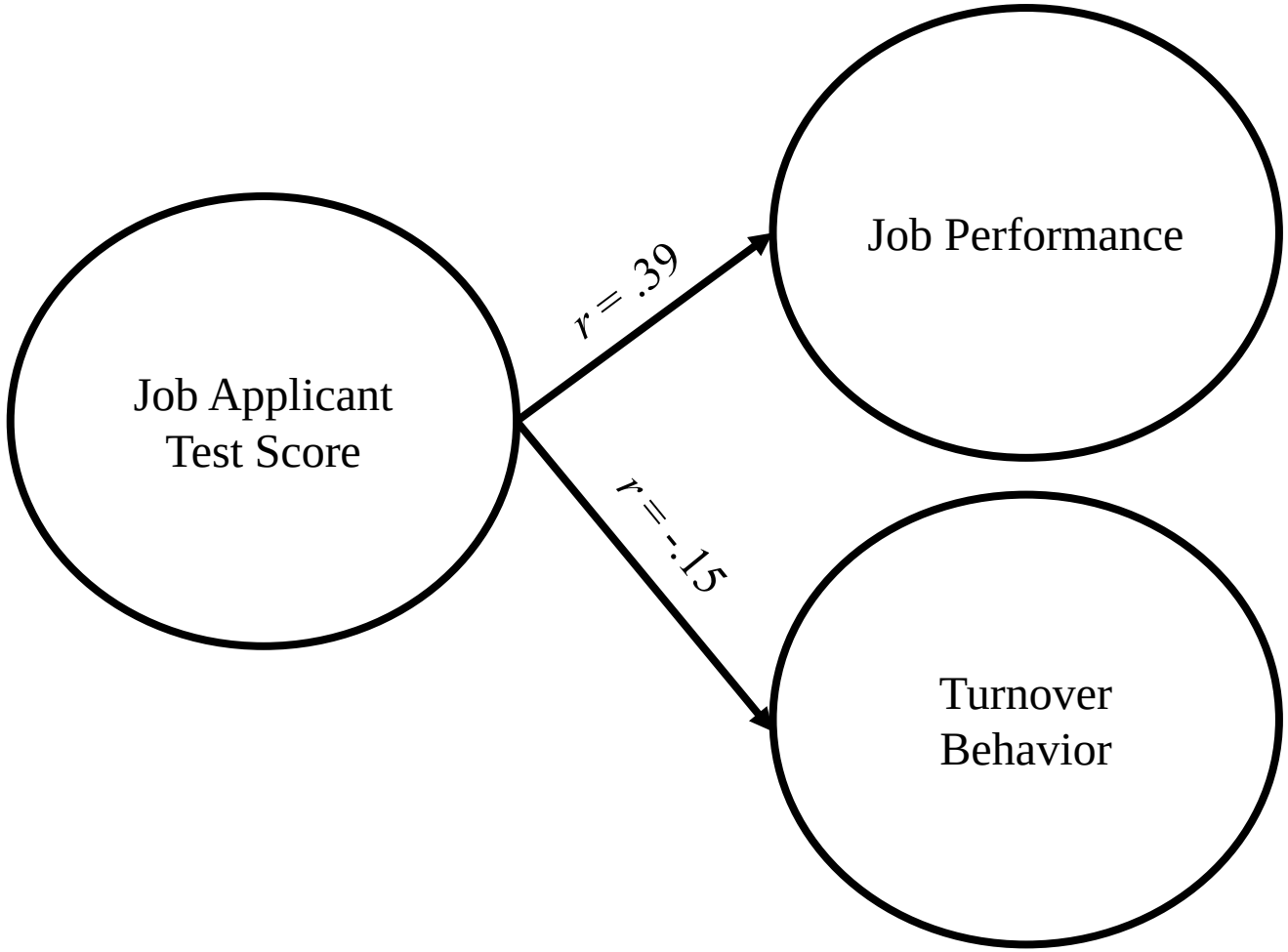
First, we can use the *correlation coefficient* to measure the **association** between variables in each of relation of interest

The correlation coefficient has two characteristics…

(1) Direction → Tells us if the nature of the relation is positive (+) or negative (-)

(2) Magnitude → Tells us if the relation between two things is "weak" or "strong"
   (1) Ranges on a scale from 0 → |1|

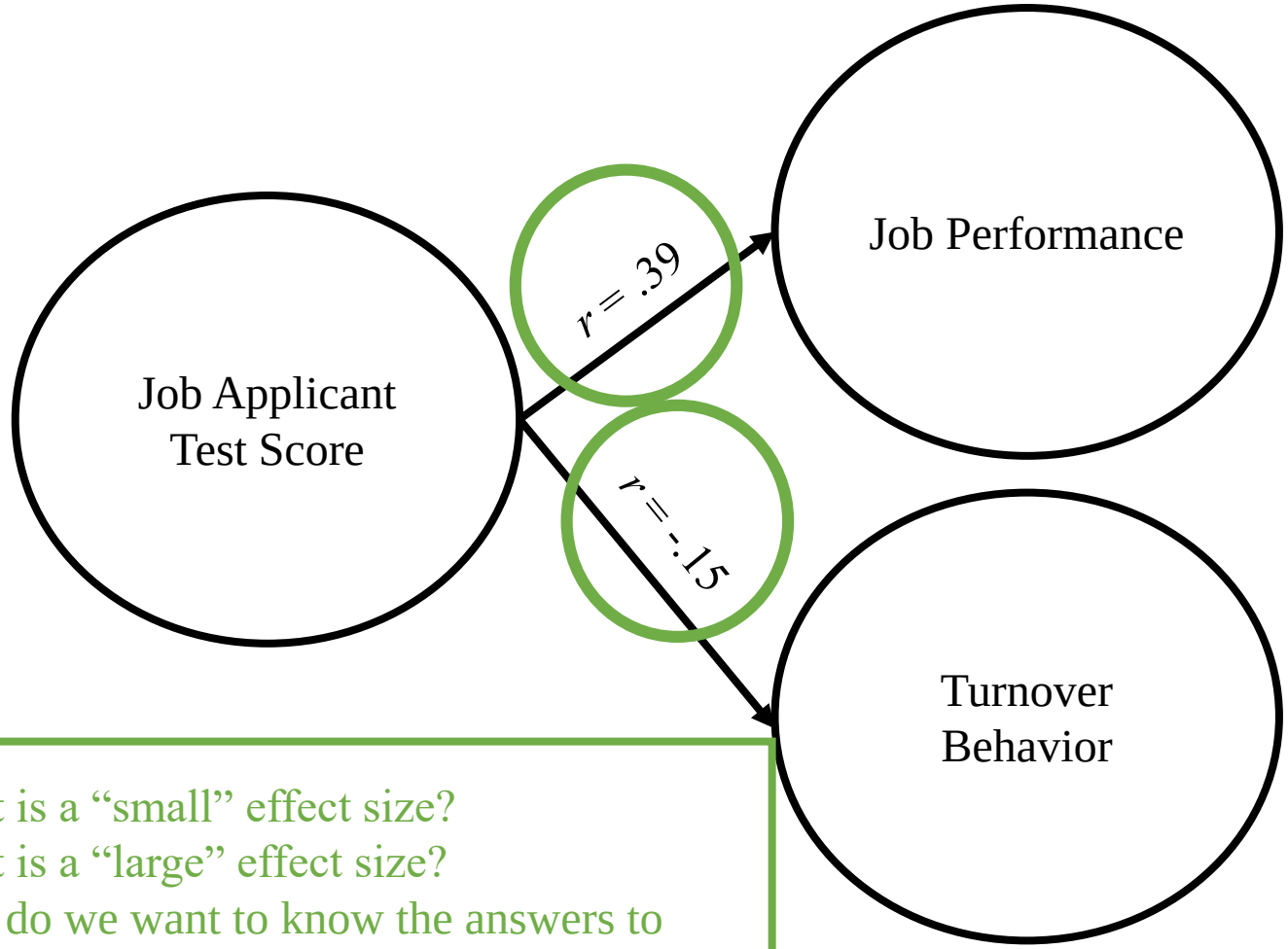
Motivating Example:



Correlation Effect Size Benchmarks

Benchmarks are standards or points of reference against which things may be compared or assessed.

Motivating Example:

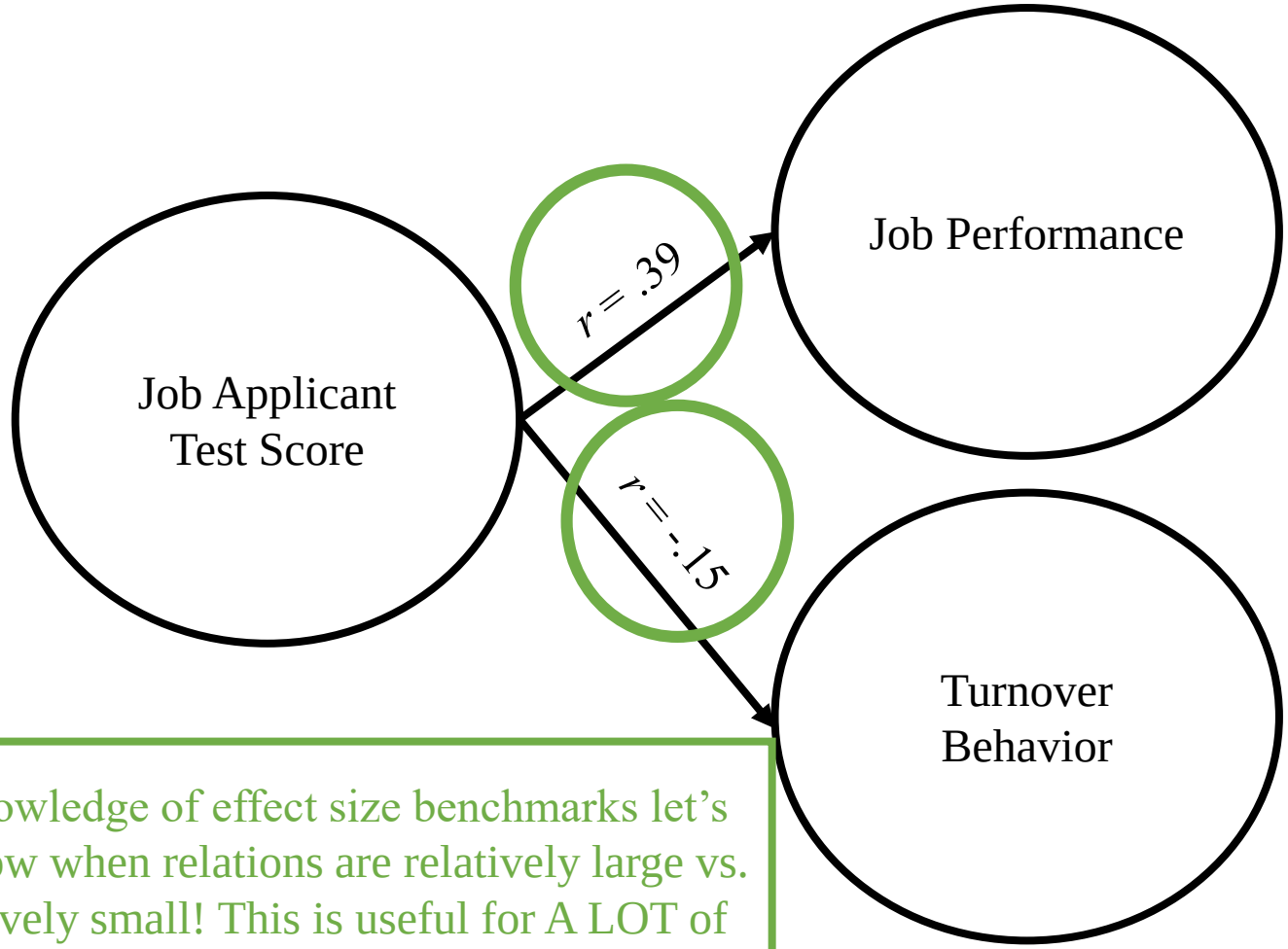


- What is a “small” effect size?
- What is a “large” effect size?
- Why do we want to know the answers to these questions?

Correlation Effect Size Benchmarks

Benchmarks are standards or points of reference against which things may be compared or assessed.

Motivating Example:

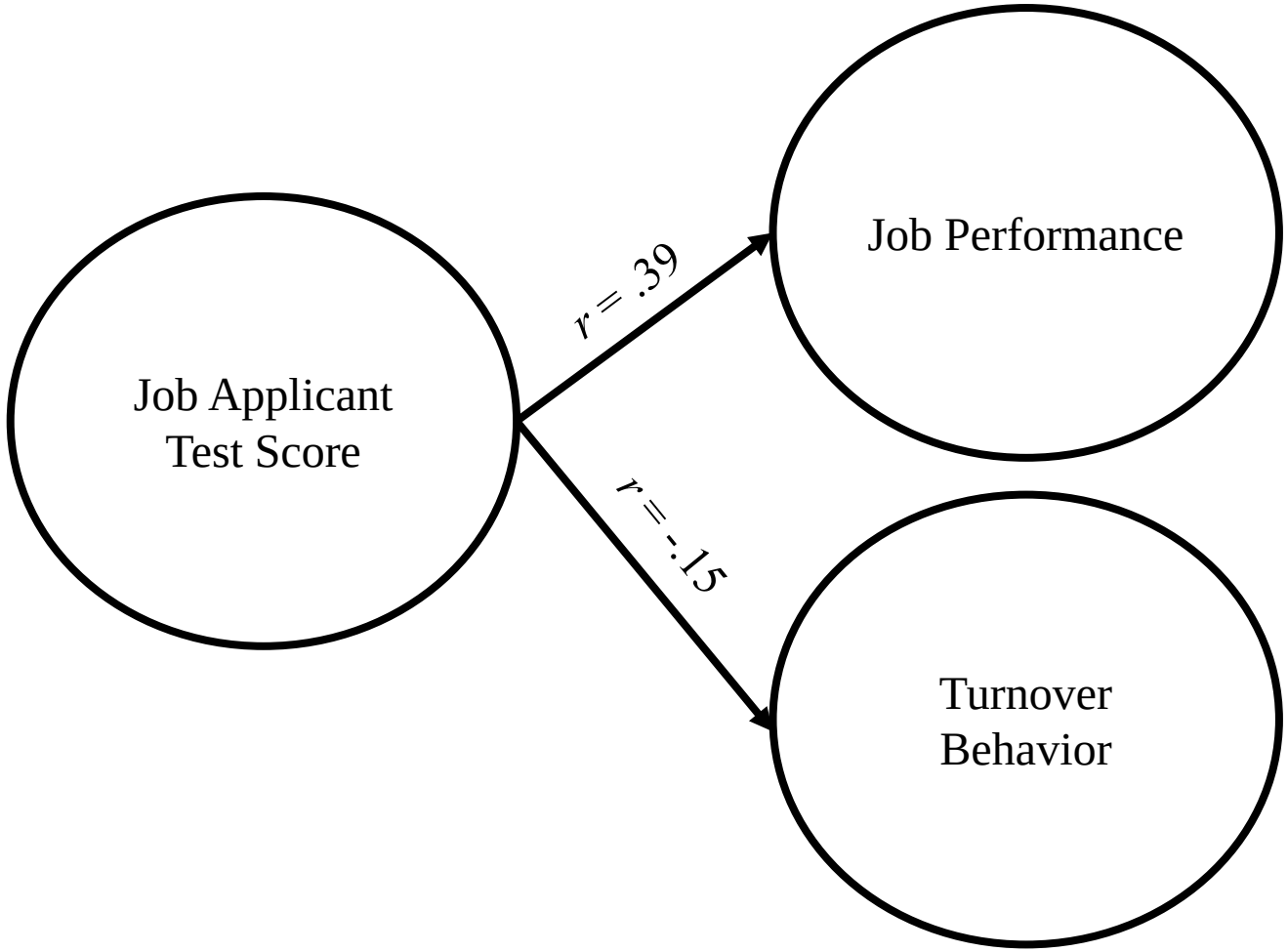


A knowledge of effect size benchmarks let's us know when relations are relatively large vs. relatively small! This is useful for A LOT of things!

Correlation Effect Size Benchmarks

Benchmarks are standards or points of reference against which things may be compared or assessed.

Motivating Example:

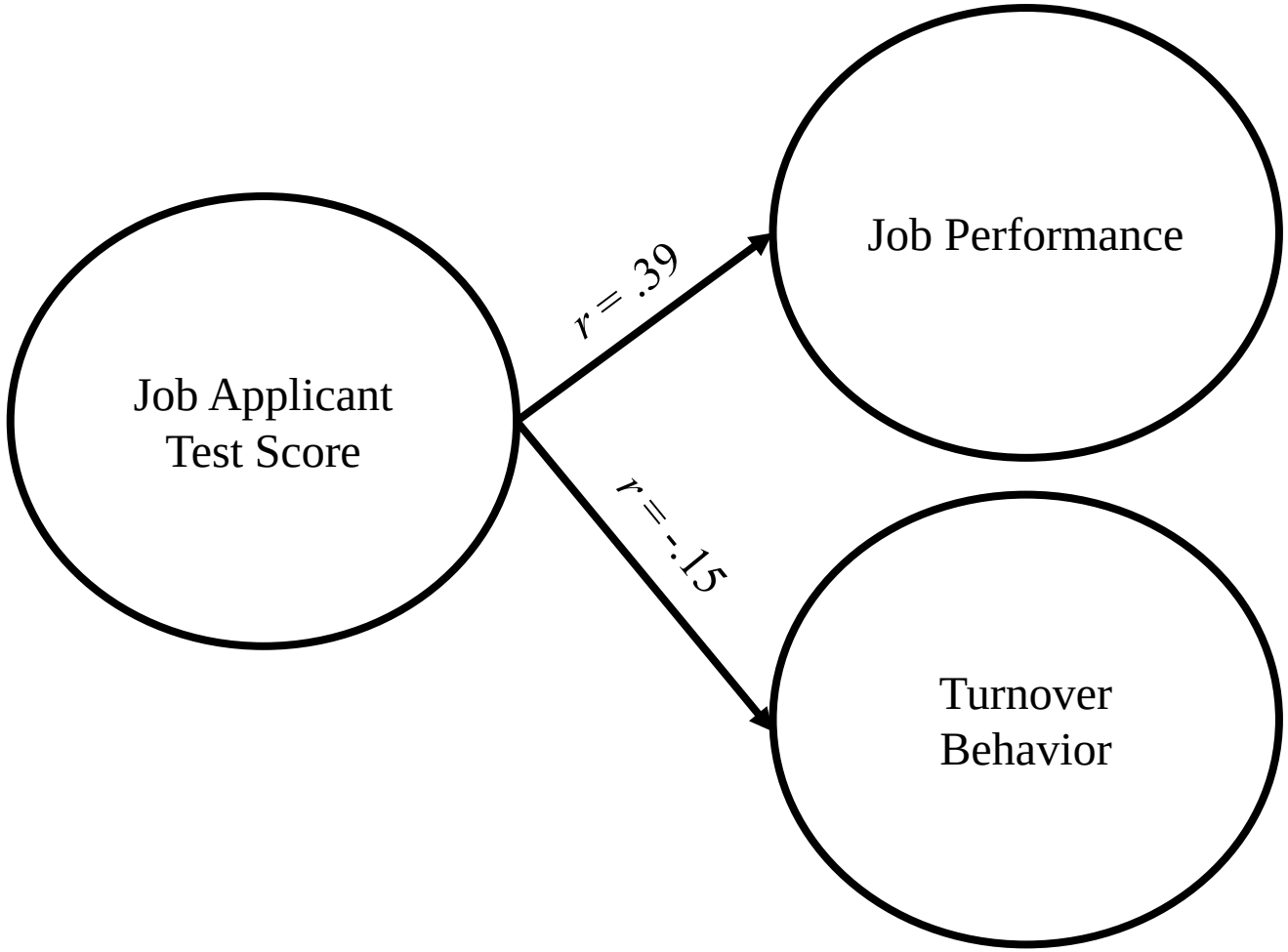


Correlation Effect Size Benchmarks

Benchmarks are standards or points of reference against which things may be compared or assessed.

With regard to correlation coefficients, benchmarks are useful because they let us know how strongly two things are associated (or correlated).

Motivating Example:



Correlation Effect Size Benchmarks

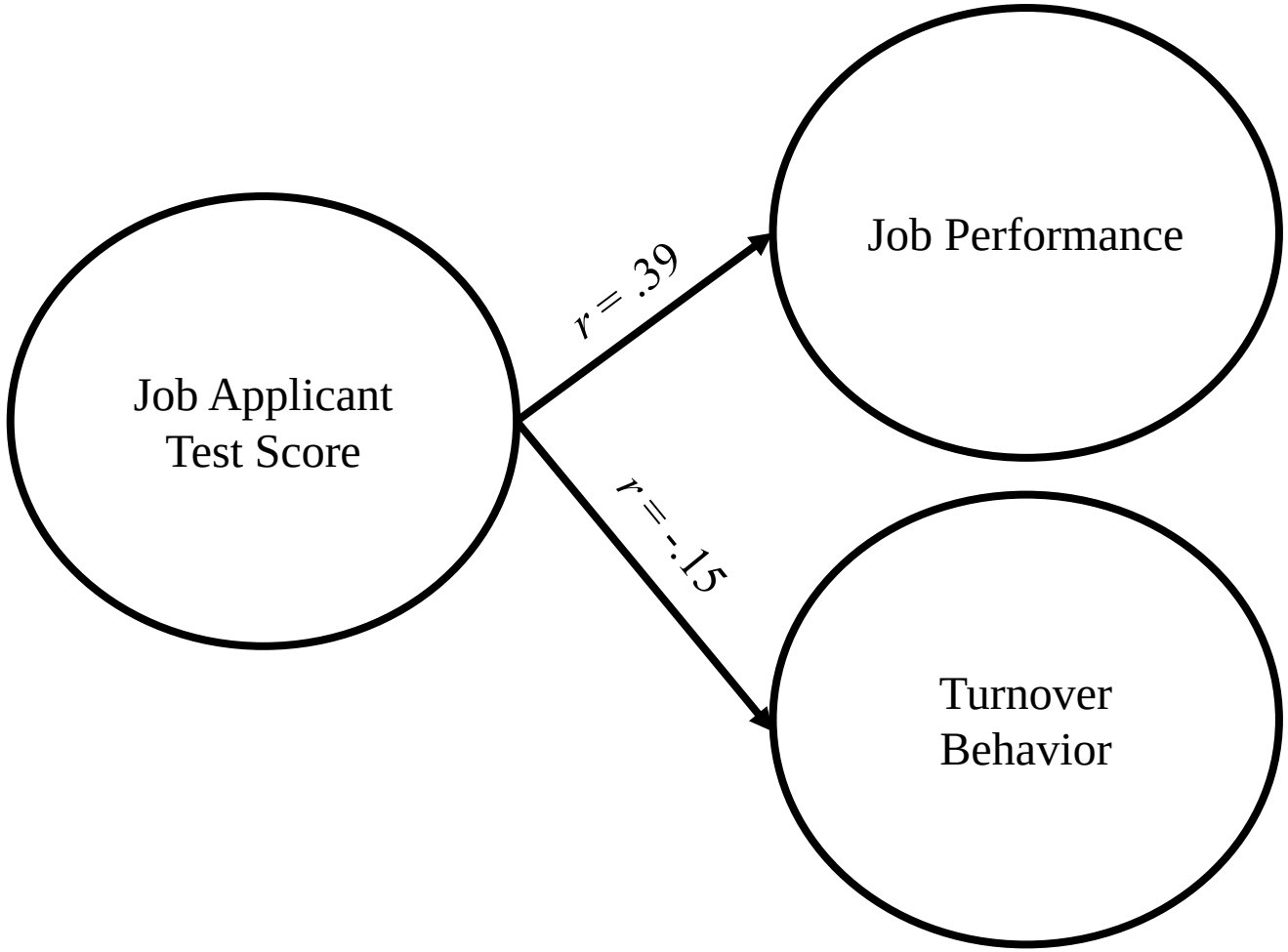
Benchmarks are standards or points of reference against which things may be compared or assessed.

With regard to correlation coefficients, benchmarks are useful because they let us know how strongly two things are associated (or correlated).

According to Cohen (1988)

- A “small” effect size = ~|.10|
- A “medium” effect size = ~|.30|
- A “large” effect size = ~|.50|

Motivating Example:



Correlation Effect Size Benchmarks

Benchmarks are standards or points of reference against which things may be compared or assessed.

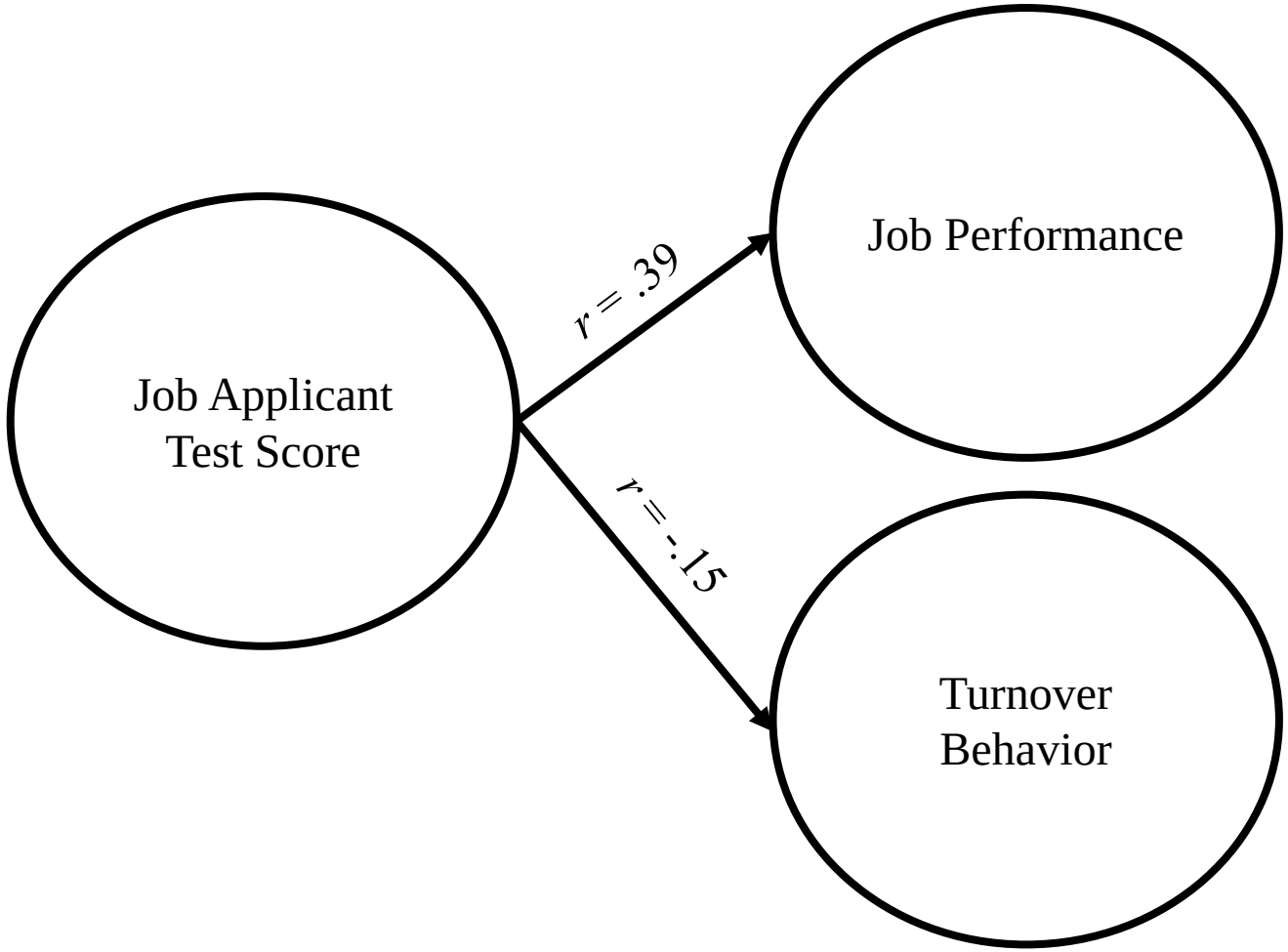
With regard to correlation coefficients, benchmarks are useful because they let us know how strongly two things are associated (or correlated).

According to Cohen (1988)

- A “small” effect size = ~|.10|
- A “medium” effect size = ~|.30|
- A “large” effect size = ~|.50|

However, these benchmarks were established arbitrarily & without evidence!

Motivating Example:



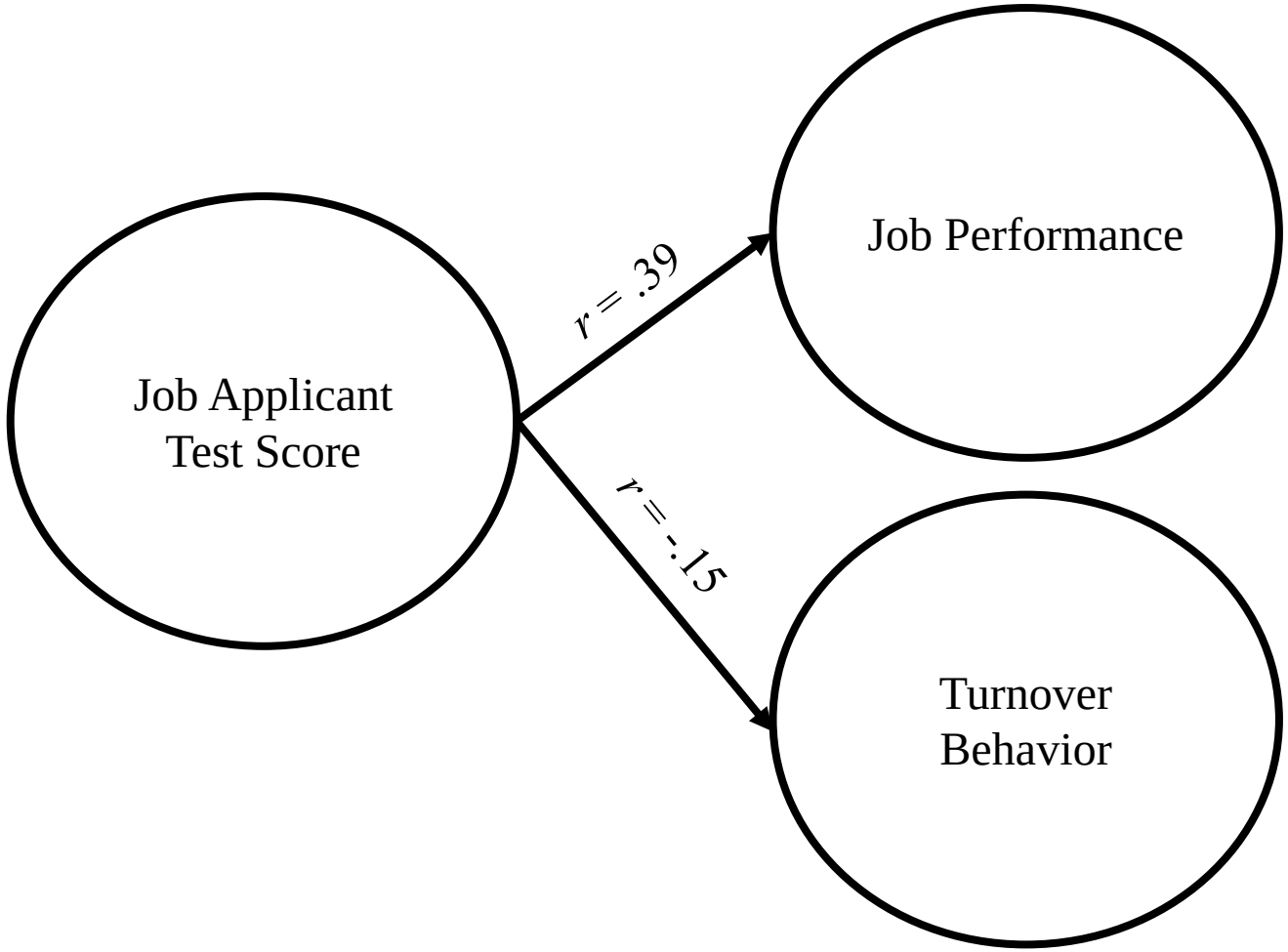
Correlation Effect Size Benchmarks

Benchmarks are standards or points of reference against which things may be compared or assessed.

Empirical evidence (see Bosco et al., 2015) suggests that effect size benchmarks should be…

- A "small" effect size = ~~|.10|~~ ~|.09|
- A "medium" effect size = ~~|.30|~~ ~|.16|
- A "large" effect size = ~~|.50|~~ ~|.32|

Motivating Example:



Correlation Effect Size Benchmarks

Benchmarks are standards or points of reference against which things may be compared or assessed.

Empirical evidence (see Bosco et al., 2015) suggests that effect size benchmarks should be…

- A "small" effect size = ~|.10| ~|.09|
- A "medium" effect size = ~|.30| ~|.16|
- A "large" effect size = ~|.50| ~|.32|

Does this affect our interpretation of the results shown in the adjacent model?

How are correlation coefficients reported?

How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

Variables included in the study

Descriptive statistics

How are correlation coefficients reported?

Intercorrelations and reliability estimates

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

Variables included in the study

Descriptive statistics

How are correlation coefficients reported?

Intercorrelations and reliability estimates

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

Variables included in the study

Descriptive statistics

How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

What is the effect size for the “relationship length→customer dependence” relation?

How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

What is the effect size for the “relationship length→customer dependence” relation?

How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

What is the effect size for the “relationship length→customer dependence” relation?

How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

What is the effect size for the “relationship length→customer dependence” relation?

How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

These are descriptive statistics, *not* correlations! So, we jump right over these!

What is the effect size for the "relationship length→customer dependence" relation?

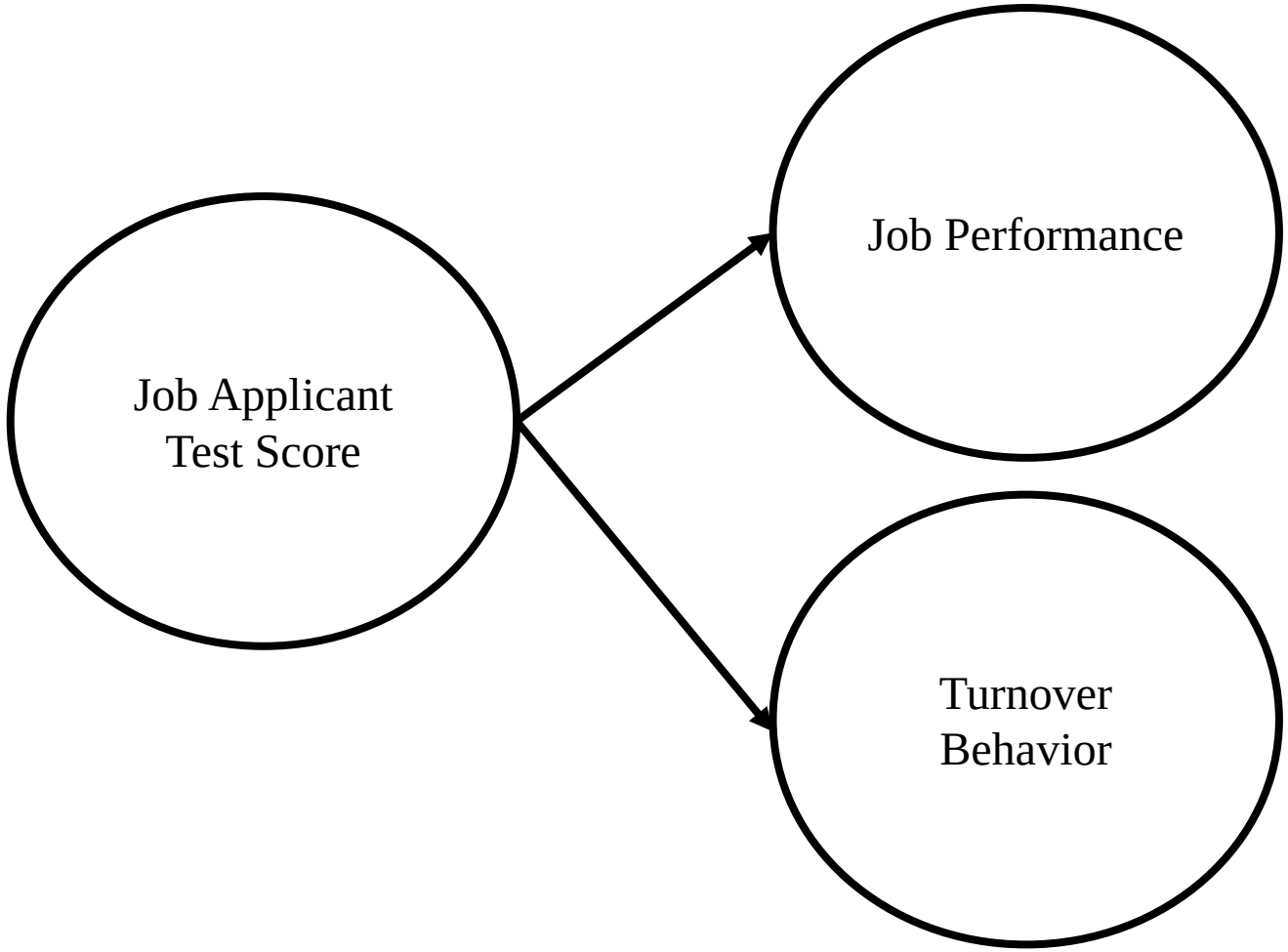
How are correlation coefficients reported?

TABLE 1. "REGULAR" CORRELATION MATRIX

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Supplier innovation	5.05	0.75	-													
2. Supplier innovation knowledge	5.43	0.99	.35**	(.83)												
3. Customer innovation know.	4.93	1.17	.29**	.29**	(.85)											
4. Embedded ties	5.58	0.86	.22**	.22**	.13	(.72)										
5. Relationship length	12.28	12.36	.03	-.03	-.04	.00	-									
6. Relationship formalization	4.28	1.49	.04	.17*	.01	.11	.02	-								
7. CRS investments	2.96	0.97	.15	.09	.15	.25**	.09	.03	(.84)							
8. Supplier financial performance	4.73	1.38	.23**	.16*	.11	.33**	.12	.02	.14	(.93)						
9. Supplier strategic advantage	5.27	1.20	.32**	.21 **	.20*	.27**	.06	-.00	.19*	.43**	(.81)					
10. Customer dependence	0.18	0.39	.07	.09	-.01	.01	-0.1	-.1	.02	.04	.03	-				
11. Market turbulence	4.30	1.18	.20*	.20*	.27**	.09	.04	.15	.13	.11	.00	-.10	(.83)			
12. Technological turbulence	4.50	1.16	.15	.14	.14	.05	.02	.19*	.11	.02	.11	.04	.40**	(.80)		
13. Opportunism	2.84	1.10	-.24**	-.26**	-.25**	-.25**	.09	.28**	-.04	-.22**	-.31**	.07	-.06	.07	(.78)	
14. Knowledge redundancy	2.94	1.26	-.17*	-.09	-.12	-.14	-.00	.12	.11	-.02	-.07	-.10	.09	.06	.07	-

What is the effect size for the “relationship length→customer dependence” relation?

Motivating Example:



In addition to the correlation coefficient, which quantifies the association between two things, one can employ a technique called ***Simple Linear Regression.***

General Linear Model

- Both correlation analysis and simple linear regression are part of a family of analysis called the *general linear model* (GLM)

- Later on, in Module 4, we will learn about multiple regression, which is another member of the GLM family
  - Simple linear regression = one predictor in the model
  - Multiple regression = multiple predictors in the model

General Linear Model

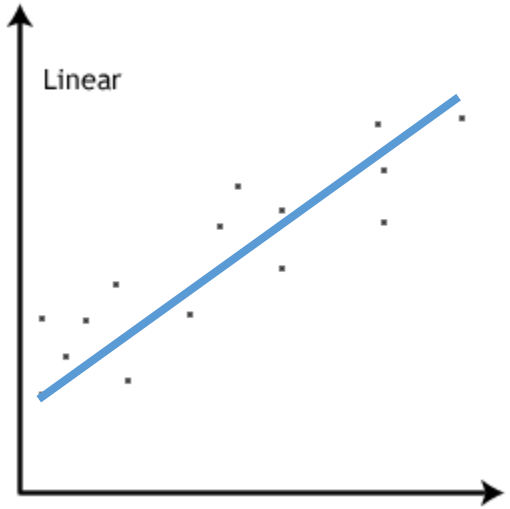
- Both correlation analysis and simple linear regression are part of a family of analysis called the *general linear model* (GLM)

- Later on, in Module 4, we will learn about multiple regression, which is another member of the GLM family
  - Simple linear regression = one predictor in the model
  - Multiple regression = multiple predictors in the model

- Although the GLM technique relies on *many* assumptions, we are only going to introduce and discuss one of them…

GLM Assumption: Linearity

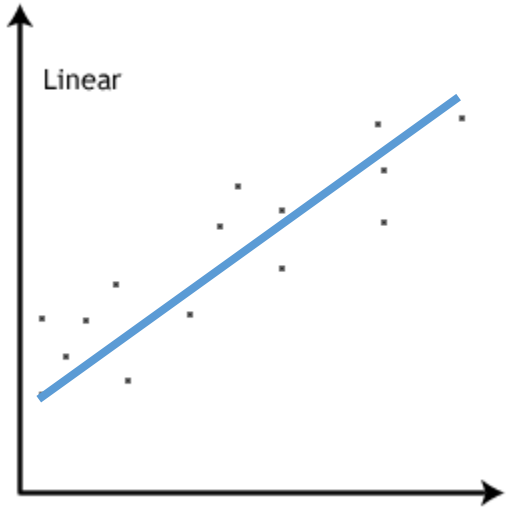
- Linearity is the assumption that the outcome variable is, in reality, linearly related to the predictor
  - Put differently, the X→Y relation can be summarized by a straight line



GLM Assumption: Linearity

- Linearity is the assumption that the outcome variable is, in reality, linearly related to the predictor
  - Put differently, the X→Y relation can be summarized by a straight line

But, will we always observe a linear relationship?

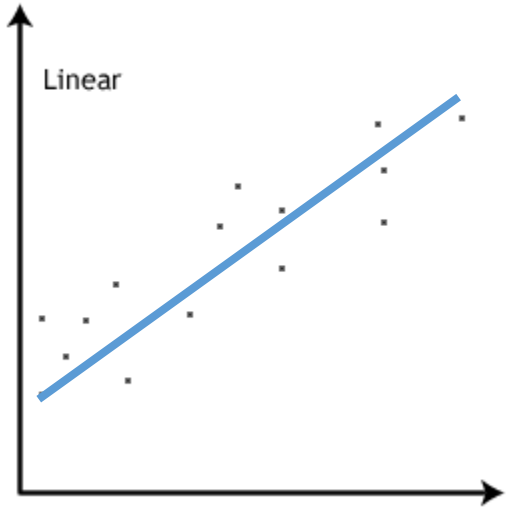


GLM Assumption: Linearity

- Linearity is the assumption that the outcome variable is, in reality, linearly related to the predictor
  - Put differently, the X→Y relation can be summarized by a straight line

But, will we always observe a linear relationship?

Can you think of examples of non-linear relationships?

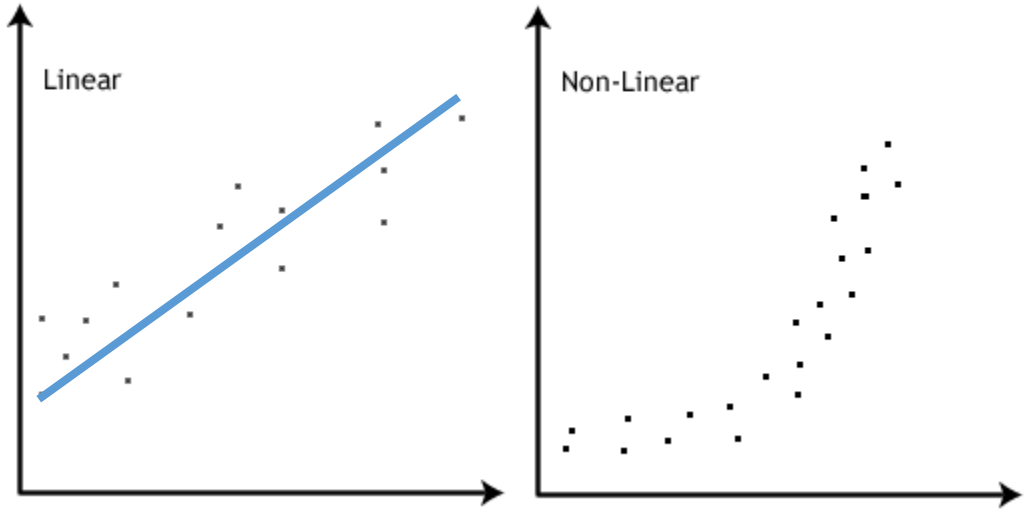


GLM Assumption: Linearity

- Linearity is the assumption that the outcome variable is, in reality, linearly related to the predictor
  - Put differently, the X→Y relation can be summarized by a straight line

But, will we always observe a linear relationship?

Can you think of examples of non-linear relationships?



GLM Assumption: Linearity

- Linearity is the assumption that the outcome variable is, in reality, linearly related to the predictor
  - Put differently, the X→Y relation can be summarized by a straight line

But, will we always observe a linear relationship?

Can you think of examples of non-linear relationships?