

Taking it again: Examining test-retest performance on video situational judgment tests for public safety candidates



NATIONAL
TESTING
NETWORK



Steve Khazon, Ph.D.
Josie Herman, M.A.



Welcome!

Agenda

- About Us
- Study Background
- Methodological Approach
- Findings
- Takeaways

About Us

- For over 40 years Ergometrics has been developing employment exams
 - Video based, human relations exams
 - Entry level and promotional
 - Fire, police, corrections, emergency communications, transit
- NTN was founded in 2006 for recruitment and exam administration
 - Job Posting
 - Exam Scheduling
 - Data Collection
 - Candidate Support
 - Reporting



- High-stakes selection tests are at least 1500 years old – retaking has been a common practice since the beginning

Why Retest?

- Fundamental to ideas of fairness and self-improvement
- Mandated by Uniform Guidelines (1978) and Standards (1999)

Benefits of Re-Takes

- Reduces impact of some random or irrelevant factors on individual performance
- Reflects changes in skills over time – People can get better!
- Builds trust and perceptions of fairness

Downsides of Re-Takes

- Increased exposure to security threats
- Associated costs adds bigger burden for less economically secure candidates
- Can be reduced or mitigated with various methods, but not always feasible

NTN's Retest Policy

- NTN has 3-month re-test policy
- Decrease the chance of compromising test content
- Waiting period supports a more competitive and inclusive process
- Tests assess situational judgment, rather than knowledge candidates can study for

Firefighter Test Components

Exam	Dimensions Measured
Video Situational Judgement Test	Considers Coworkers Before Self
	Congenial with Coworkers
	Leadership
	Industrious/Hardworking
	Adaptability
	Customer Relations
	Communication
	Compassion
	Dealing with Difficult Situations
	Integrity/Ethics
	Supervisory/Management Support
	Persistence in Ongoing Learning
Mechanical Reasoning	Mechanical Reasoning, Understanding Systems
Math	Math
Reading	Reading
Public Safety Self Assessment	Integrity
	Use of Authority
	Commitment to Equality
	Group Bias

Impact of Multiple Retakes

- Common part of test life-cycle
- Scores generally increase across re-tests (Hausknecht et al., 2007)
- *Q1: How do NTN candidates' test scores change across multiple retakes of the same exam?*

Impact of Waiting to Retakes

- Candidates motivated to re-test quickly
- Waiting provides time to improve relevant skills, or forget them
- *Q2: How do NTN candidate's test scores change as a function of the time a candidate waited to retake the same exam?*

Impact of Demographic Factors

- Re-testing confers a resource burden that better-off candidates can more easily bear (when candidates shoulder cost)
- *Q3: What is the impact of socio-demographic variables on NTN test score change?*



NATIONAL
TESTING
NETWORK

Methodological Approach

Participants

- NTN fire entry-level candidates testing 2018 through 2022
- Total N = 72,889
- Represents every US state and territory except for American Samoa

Sample as Percent of Total N

	N	%
Sex		
Female	6892	9.46%
Male	64902	89.04%
Undisclosed	1095	1.50%
Ethnicity		
American Indian or Alaskan Native	733	1.01%
Asian	2759	3.79%
Black or African American	5202	7.14%
Hispanic or Latino	13336	18.30%
Native Hawaiian or Other Pacific Islander	965	1.32%
Undisclosed	3633	4.98%
White	46261	63.47%
Total	72889	100.00%

Measures

- Entry-Level Fire Service Video Situational Judgement Test (Swander et al., 2021)
- Socio-Demographic Info (United States Census Bureau, 2020)
 - Regional income inequality
 - Regional poverty
 - Major urban area



NATIONAL
TESTING
NETWORK

Findings

Entry-Level Fire

Difference			
Score	N	Mean	Std Dev
T2-T1	23678	2.37	8.09
T3-T2	8940	2.01	7.64
T4-T3	3537	1.85	7.34
T5-T4	1399	1.69	7.14
T6-T5	552	1.34	6.33
T7-T6	224	1.32	6.23
T7-T8	94	1.28	6.34

- *Q1: How do NTN candidates' test scores change across multiple retakes of the same exam?*
- Small increase across time, but gain is smaller on subsequent retakes
- No benefit after 7 retakes

Entry-Level Scores Across Test Sessions

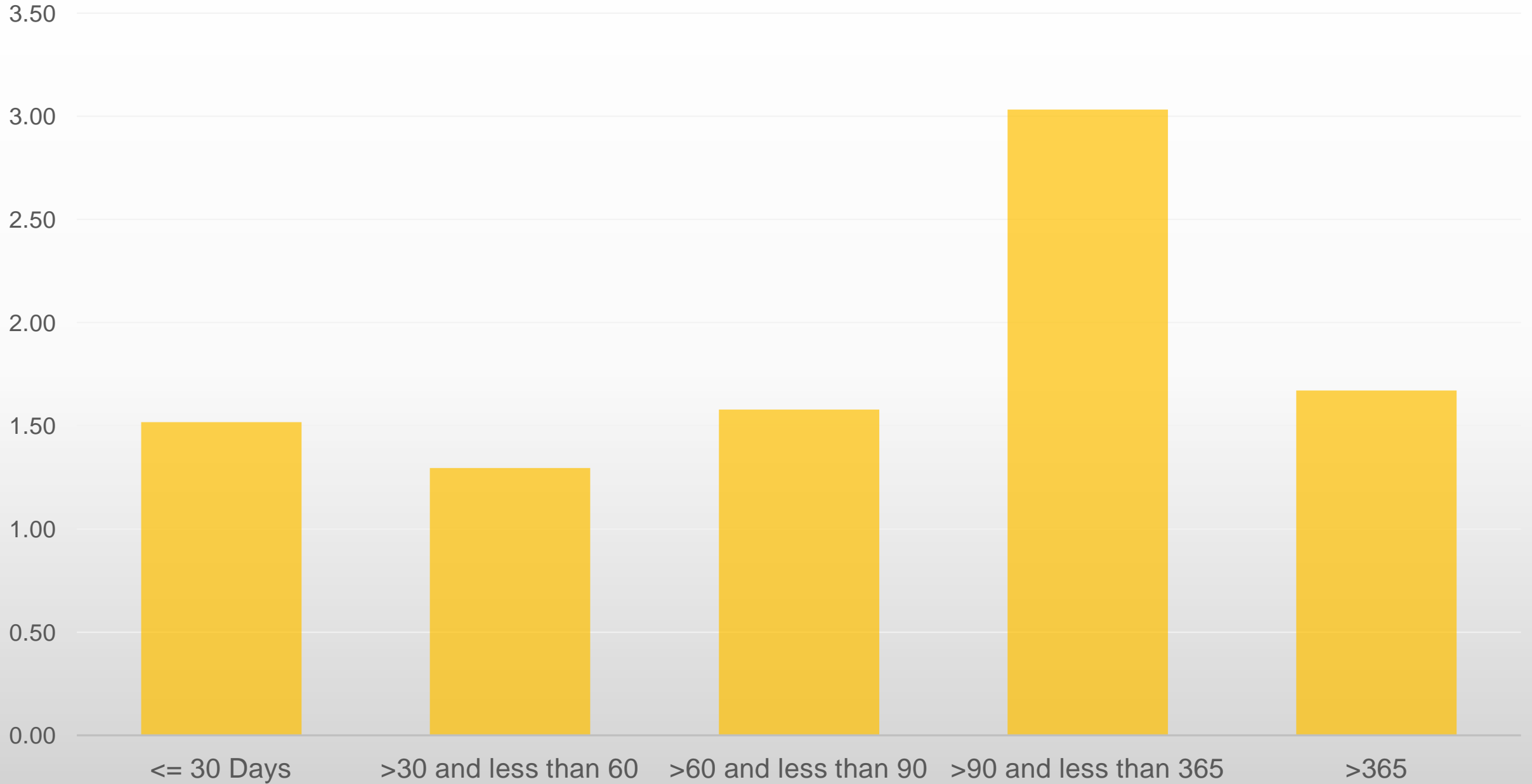


Entry-Level Fire

Time Between Retakes		N	Mean	Std. Deviation
None	T1	49211	82.21	8.71
	T2	0		
	T1/T2 Diff	0		
<= 30 Days	T1	962	82.51	9.51
	T2	962	84.03	9.15
	T1/T2 Diff	962	1.52	8.01
>30 and less than 60	T1	630	82.77	7.91
	T2	630	84.06	8.96
	T1/T2 Diff	630	1.30	7.36
>60 and less than 90	T1	567	82.54	7.91
	T2	567	84.12	8.60
	T1/T2 Diff	567	1.58	8.25
>90 and less than 365	T1	12533	80.66	8.73
	T2	12533	83.69	8.85
	T1/T2 Diff	12533	3.03	8.36
>365	T1	8986	83.10	8.07
	T2	8986	84.77	8.26
	T1/T2 Diff	8986	1.67	7.66

- *Q2: How do NTN candidate's test scores change as a function of the time a candidate waited to retake the same exam?*
- Most benefit after waiting between 3 months and 1 year

Entry-Level T1-T2 Differences by Time Between Retakes



Variable	Standardized Coefficients	t-score	Sig.
(Constant)		3.54	0.00
Regional Income Inequality	0.01	1.07	0.29
Lives in Major Urban Area	0.00	0.53	0.60
% Own Race Below Poverty	-0.01	-0.90	0.37
% Own Sex Below Poverty	0.01	0.75	0.45
Ethnic Minority Status	0.00	-0.01	0.99
Female Candidate Status	0.00	0.58	0.56

Outcome: Difference score (T2 – T1)

- *Q3: What is the impact of socio-demographic variables on NTN test score change?*
- Regional inequality approaches significance



NATIONAL
TESTING
NETWORK

Takeaways

Summary of Findings

- Differences between scores on retake and across time generally small – largest effects after 1 or 2 retakes
- Largest score boost for retakes from 3 months to a year
- No significant effects from socio-demographic factors on retake score boost

Takeaways

- Video SJTs are resilient to learning effects
- Enforcing a waiting period for candidates and explaining why is recommended
- Providing support to candidates in high-inequality areas may ameliorate differences
- Larger increase overtime may be due to candidate's personal growth and increased exposure to life events



Thank You!

Please visit us at:

www.nationaltestingnetwork.com