Impact of Education and Experience Level on the Effectiveness of Exploratory Testing: An Industrial Case Study

Ceren Şahin Gebizli
Vestel Electronics, R&D
ceren.sahin@vestel.com.tr

Hasan Sözer
Özyeğin University
hasan.sozer@ozyegin.edu.tr

ICSTW TAIC-PART 2017: 12th Workshop on Testing: Academic and Industrial Conference – Practice and Research Techniques (TAIC PART)

March 13th, 2017, Tokyo, Japan

friendly technology VESTEL

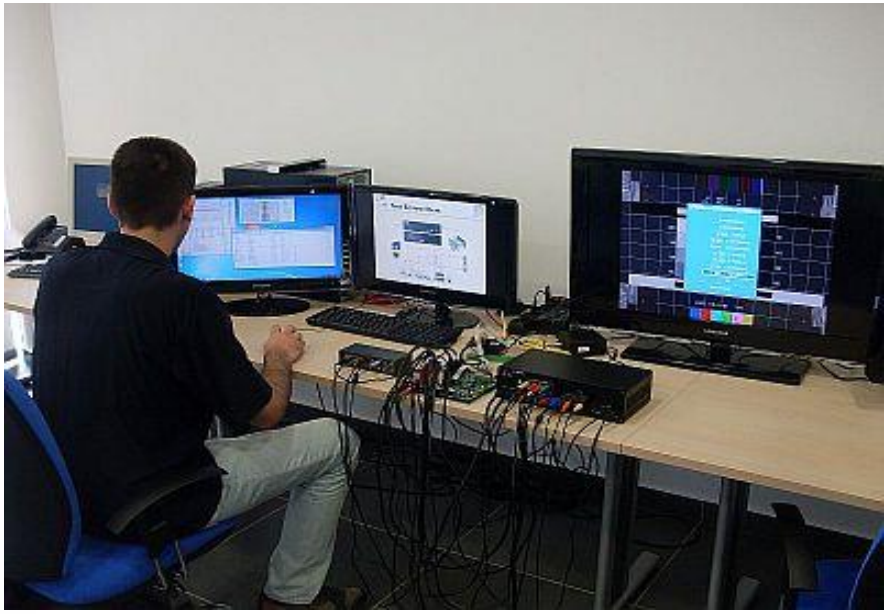
ÖZYEĞİN UNIVERSITY
SOFTWARE RESEARCH LAB

Setting the Context: Consumer Electronics Domain

- 900+ Customers
  - 157 Different Brands
  - 145 Countries
- 100 Software Engineers
- 100 Test Engineers/ Technicians
- 10M DTV production annually
- Short time-to-market
- Test Effectiveness is important

Exploratory Testing (ET) Approach

- Test engineers/ technicians perform manual tests
- Iterative Process
  - Learn about the product;
  - Plan the testing work to be done;
  - Design and execute the tests;
  - Report the results.



Motivation

- ET proved effective in detecting critical failures
- Manual task; hence, assumed to be dependent on background and experience
- Lack of evaluation in industrial context
- Lack of empirical studies
- Goal: Evaluate the impact of the educational backgrounds and experience levels of testers on the effectiveness of ET
- Context: Testing Smart TV systems developed by Vestel

Research Questions

- **RQ1**: How domain and testing experiences are affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ2:** How domain and testing experiences are affecting the number of critical failures detected?
- **RQ3:** How educational background is affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ4:** How educational background is affecting the number of critical failures detected?

Experimental Setup

- Participant properties
  - Domain Experience
  - Testing Experience
  - Higher Education
- Collected metrics
  - Test duration
  - Number of failures detected
  - Efficiency: Number of failures detected / Test duration

List of Participants

Practitioner ID (PId)	Domain Experience (# of years)	Testing Experience (# of years)	Higher Education (Yes/No)
1	11	11	No
2	8	8	No
3	16	14	No
4	7	5	No
5	7	7	No
6	11	10	No
7	8	10	No
8	1	1	Yes
9	6	6	Yes
10	1	8	Yes
11	1	1	Yes
12	1	1	Yes
13	1	1	Yes
14	1	1	Yes
15	4	4	Yes
16	4	4	Yes
17	12	12	Yes
18	0	3	Yes
19	0	2	Yes

Overall Results

Practitioner ID (PId)	Test Duration (# of days)	# of Failures Detected			
		Critical	Major	Minor	Trivial
1	10	2	2	5	1
2	10	2	2	6	2
3	9	2	2	5	0
4	9	2	2	5	0
5	10	2	2	4	0
6	10	2	2	5	0
7	10	2	2	5	0
8	15	2	2	5	2
9	6	2	2	6	0
10	10	1	2	2	0
11	15	1	2	5	1
12	14	1	2	5	0
13	15	1	2	5	1
14	12	0	2	5	1
15	8	2	2	6	0
16	9	2	2	6	0
17	6	2	2	6	0
18	16	1	2	4	2
19	16	1	2	4	2

Research Questions

- **RQ1**: How domain and testing experiences are affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ2:** How domain and testing experiences are affecting the number of critical failures detected?
- **RQ3:** How educational background is affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ4:** How educational background is affecting the number of critical failures detected?

Impact of domain and testing experience

- Participants, who do not have higher education

Practitioner ID (PId)	Domain Experience	Testing Experience	#Critical Failures	Efficiency
1	11	11	2	1.00
2	8	8	2	1.20
3	16	14	2	1.00
4	7	5	2	1.00
5	7	7	2	0.80
6	11	10	2	0.90
7	8	10	2	0.90

Impact of domain and testing experience

- Participants, who have higher education

Practitioner ID (PId)	Domain Experience	Testing Experience	#Critical Failures	Efficiency
8	1	1	2	0.73
9	6	6	2	1.67
10	1	8	1	0.50
11	1	1	1	0.60
12	1	1	1	0.57
13	1	1	1	0.60
14	1	1	0	0.67
15	4	4	2	1.25
16	4	4	2	1.11
17	12	12	2	1.67
18	0	3	1	0.56
19	0	2	1	0.56

Impact of domain and testing experience

- **Group A**; consists of experienced subjects (who have 2 or more years of experience),
- **Group B**; consists of inexperienced subjects (who have less than 2 years of experience).

- **T-test** suggests significant difference among the groups
  - P-values << 0.05

Research Questions

- **RQ1**: How domain and testing experiences are affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ2: How domain and testing experiences are affecting the number of critical failures detected?**
- **RQ3:** How educational background is affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ4:** How educational background is affecting the number of critical failures detected?

Impact of domain and testing experience on the Criticality of Detected Failures

Practitioner ID (PId)	Domain Experience	Testing Experience	#Critical Failures	Efficiency
8	1	1	2	0.73
9	6	6	2	1.67
10	1	8	1	0.50
11	1	1	1	0.60
12	1	1	1	0.57
13	1	1	1	0.60
14	1	1	0	0.67
15	4	4	2	1.25
16	4	4	2	1.11
17	12	12	2	1.67
18	0	3	1	0.56
19	0	2	1	0.56

- **T-test** suggests significant impact for the same grouping of participants
  - P-values << 0.05

Research Questions

- **RQ1**: How domain and testing experiences are affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ2:** How domain and testing experiences are affecting the number of critical failures detected?
- **RQ3:** How educational background is affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ4:** How educational background is affecting the number of critical failures detected?

Impact of Higher Education

- Results for participants who have at least 2 years of experience

Practitioner ID (PId)	Higher Education	# of Critical Failures	Efficiency
1	No	2	1.00
2	No	2	1.20
3	No	2	1.00
4	No	2	1.00
5	No	2	0.80
6	No	2	0.90
7	No	2	0.90
9	Yes	2	1.67
15	Yes	2	1.25
16	Yes	2	1.11
17	Yes	2	1.67

Impact of Higher Education

- Results for subjects who have less than 2 years of experience

Practitioner ID (PId)	Higher Education	# of Critical Failures	Efficiency
8	Yes	2	0.73
10	Yes	1	0.50
11	Yes	1	0.60
12	Yes	1	0.57
13	Yes	1	0.60
14	Yes	0	0.67
18	Yes	1	0.56
19	Yes	1	0.56

Impact of Higher Education

- **Group C**; consists of subjects with higher education
- **Group D**; consists of subjects without higher education

- **T-test** suggests significant difference among the groups
  - P-values << 0.05

Research Questions

- **RQ1**: How domain and testing experiences are affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ2:** How domain and testing experiences are affecting the number of critical failures detected?
- **RQ3:** How educational background is affecting the test efficiency in terms of number of failures detected per unit of time?
- **RQ4:** How educational background is affecting the number of critical failures detected?

Impact of domain and testing experience on the Criticality of Detected Failures

No impact of higher education observed

Practitioner ID (PId)	Higher Education	# of Critical Failures	Efficiency
1	No	2	1.00
2	No	2	1.20
3	No	2	1.00
4	No	2	1.00
5	No	2	0.80
6	No	2	0.90
7	No	2	0.90
9	Yes	2	1.67
15	Yes	2	1.25
16	Yes	2	1.11
17	Yes	2	1.67

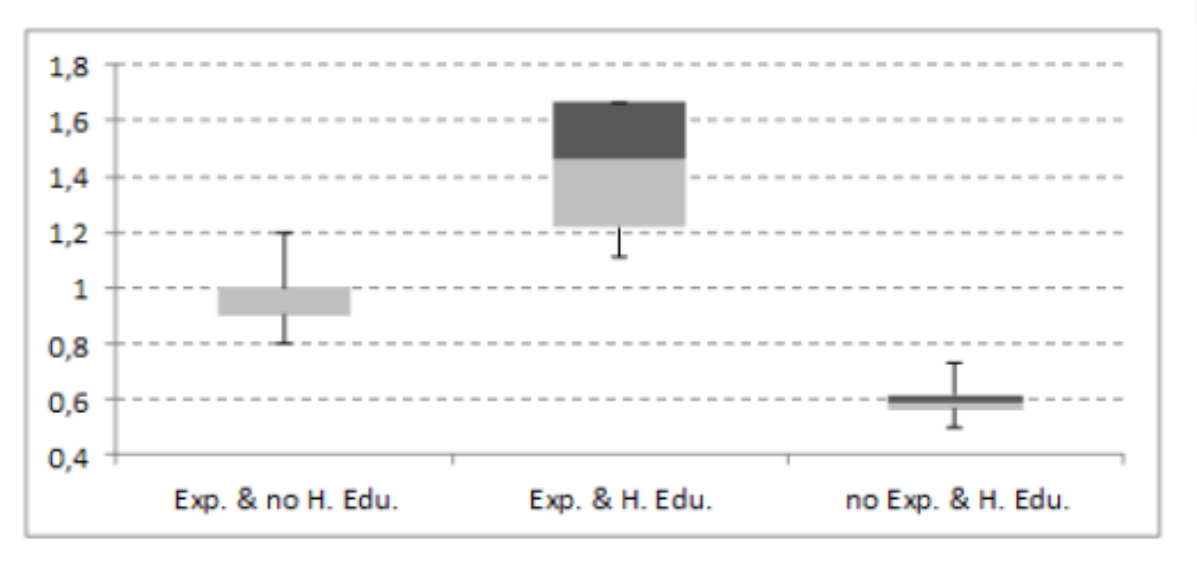
ANOVA Analysis

- Suggests significant difference among the groups of subjects
  - P-value << 0.001

- **Group A**; consists of experienced subjects (who have 2 or more years of experience),
- **Group B**; consists of inexperienced subjects (who have less than 2 years of experience).
- **Group C**; consists of subjects with higher education
- **Group D;** consists of subjects without higher education

Factors	Higher Education	No Higher Education
Experience	Members of both Group A and C	Members of both Group A and D
No Experience	Members of both Group B and C	Members of both Group B and D

Low Variance within the Groups



- Box plot regarding the test efficiency of the 3 groups; Exp. & no H. Edu (Group A & D), Exp. & H. Edu. (Group A & C), no Exp. & H. Edu. (Group B & C)

Conclusions

- Evaluating the **impact of education** level and **experience** level of testers on the effectiveness of **Exploratory Testing**
- **Case study** with 19 practitioners
- **Industrial Context**: consumer electronics domain (Smart TVs)

- Both the **educational background** and **experience** have **significant impact** on **test efficiency**
- **Experience** level has also a **significant impact** on the number of detected **critical failures**, **education level has not**

Academic-Industrial Collaboration

- Ph.D. student at Ozyegin University
  and Test Architect at Vestel Electronics R&D
- University & Company collaboration for 6 years
- Conference papers, journal articles, joint grant of Vestel Electronics and the Turkish Ministry of Science, Industry and Technology (909.STZ.2015).

Sozer, H. & Gebizli, C. S.
Model-Based Testing of Digital TVs: An Industry-as-Laboratory Approach
Software Quality Journal, 2016
DOI: 10.1007/s11219-016-9321-y

ÖZYEĞİN UNIVERSITY
SOFTWARE RESEARCH LAB