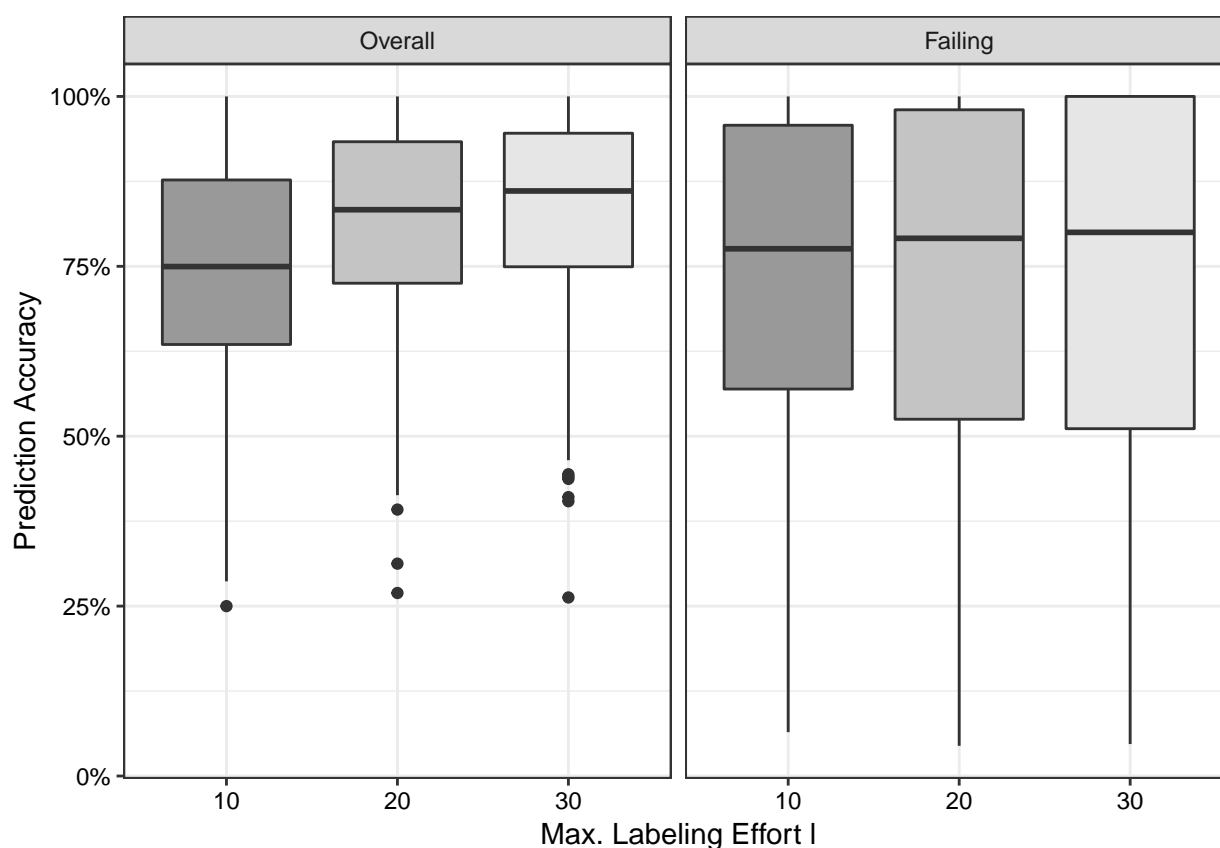# ICST

```
library(ggplot2)
library(scales)
theme_set(theme_bw())
```

## Oracle Quality

```
## [1] "Subjects (-l 10):  552"
```

```
## [1] "Subjects (-l 20):  545"
```

```
## [1] "Subjects (-l 30):  544"
```
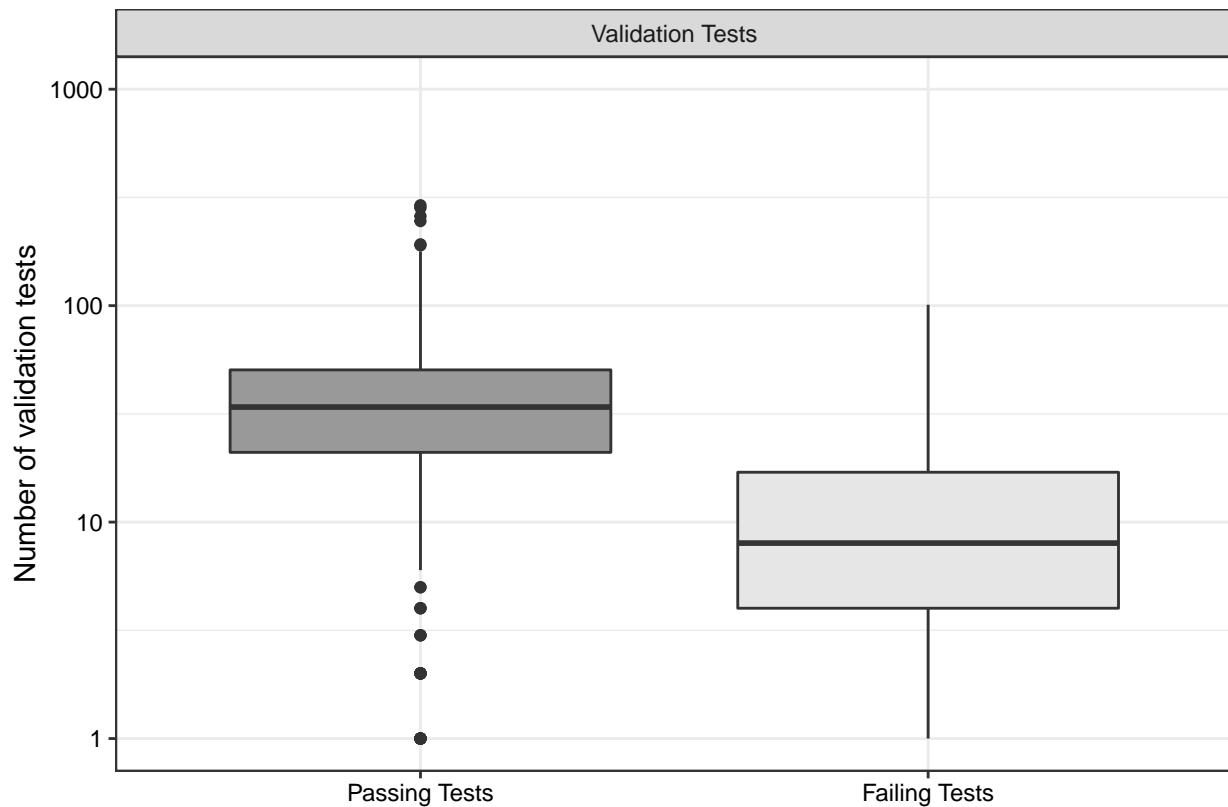


```
## Saving 5.6 x 3.6 in image
```
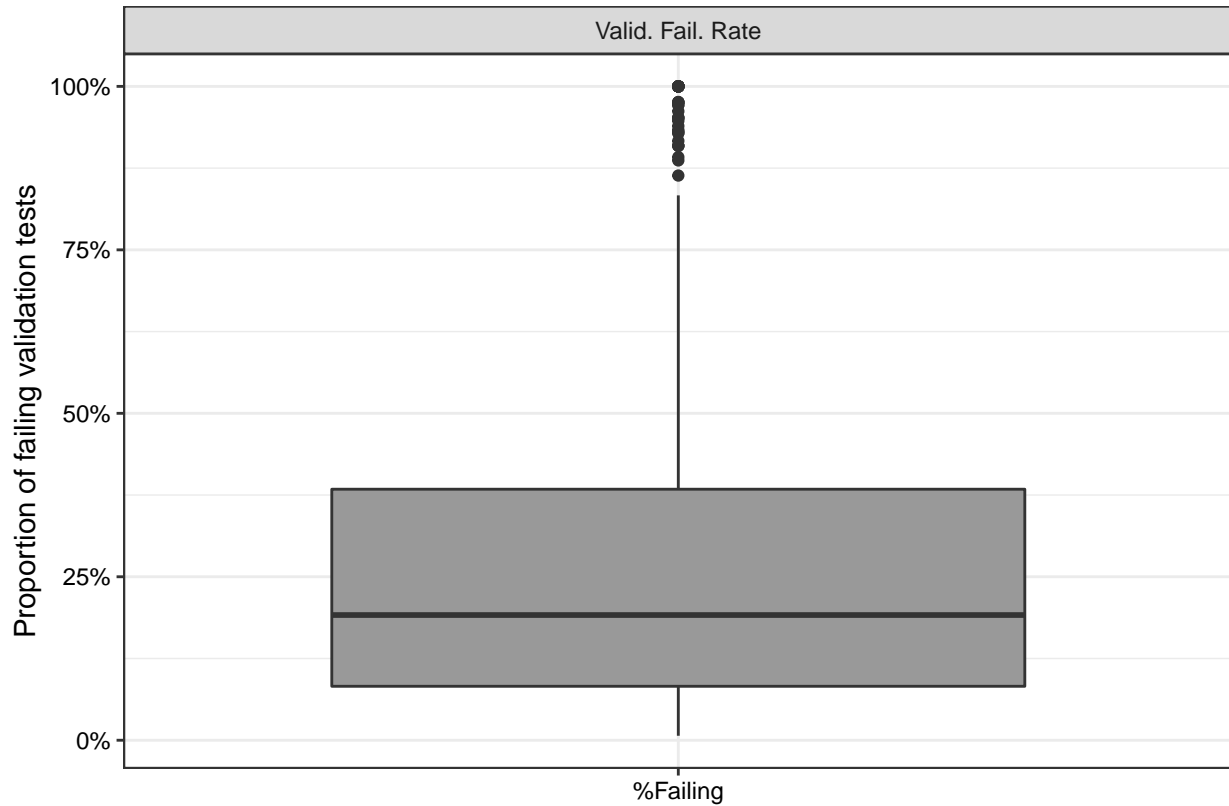
```
##                          subject max_label category variable      value
## 1     1-A-bug-18353198-18353306         10 Accuracy  Overall 0.8344828
## 3   104-A-bug-11691772-11691810         10 Accuracy  Overall 0.9149425
## 5   104-A-bug-11691779-11691810         10 Accuracy  Overall 0.9722222
## 7   104-A-bug-11691794-11691810         10 Accuracy  Overall 0.6666667
## 9   104-A-bug-11939068-11939085         10 Accuracy  Overall 0.9560284
## 11 104-A-bug-13139540-13139570         10 Accuracy  Overall 0.6864583
```

```
##        0%       25%       50%       75%      100%
## 0.2500000 0.6349879 0.7496564 0.8771287 1.0000000
```

```
## [1] "Average/Median prediction accuracy (-l 10):  0.753890193684369 0.749656419529837"

## [1] "Average/Median prediction accuracy (-l 20):  0.815275237281484 0.833333333333333"

## [1] "Average/Median prediction accuracy (-l 30):  0.838733130775872 0.860903532608696"

## [1] "Average/Median conditional accuracy (-l 10):  0.739569856633927 0.775833333333333"

## [1] "Average/Median conditional accuracy (-l 20):  0.72899397780596 0.791228070175439"

## [1] "Average/Median conditional accuracy (-l 30):  0.73405715729944 0.8"

## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Removed 17 rows containing non-finite values (stat_boxplot).
```
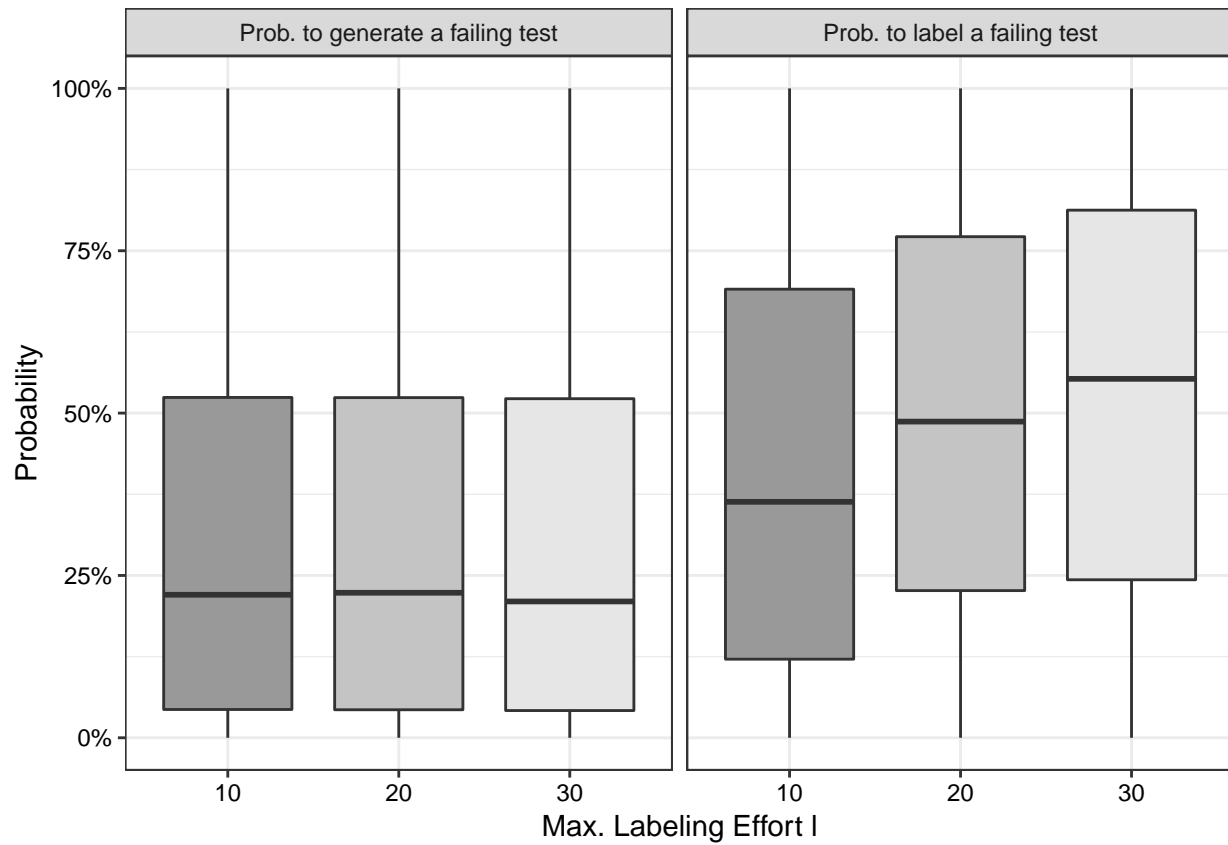


```
## Saving 2.4 x 3.6 in image

## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Removed 17 rows containing non-finite values (stat_boxplot).

## [1] "Average number of passing tests:  41.5087285902503"

## [1] "Average number of failing tests:  13.0419960474308"
```

Valid. Fail. Rate

Proportion of failing validation tests

100%

75%

50%

25%

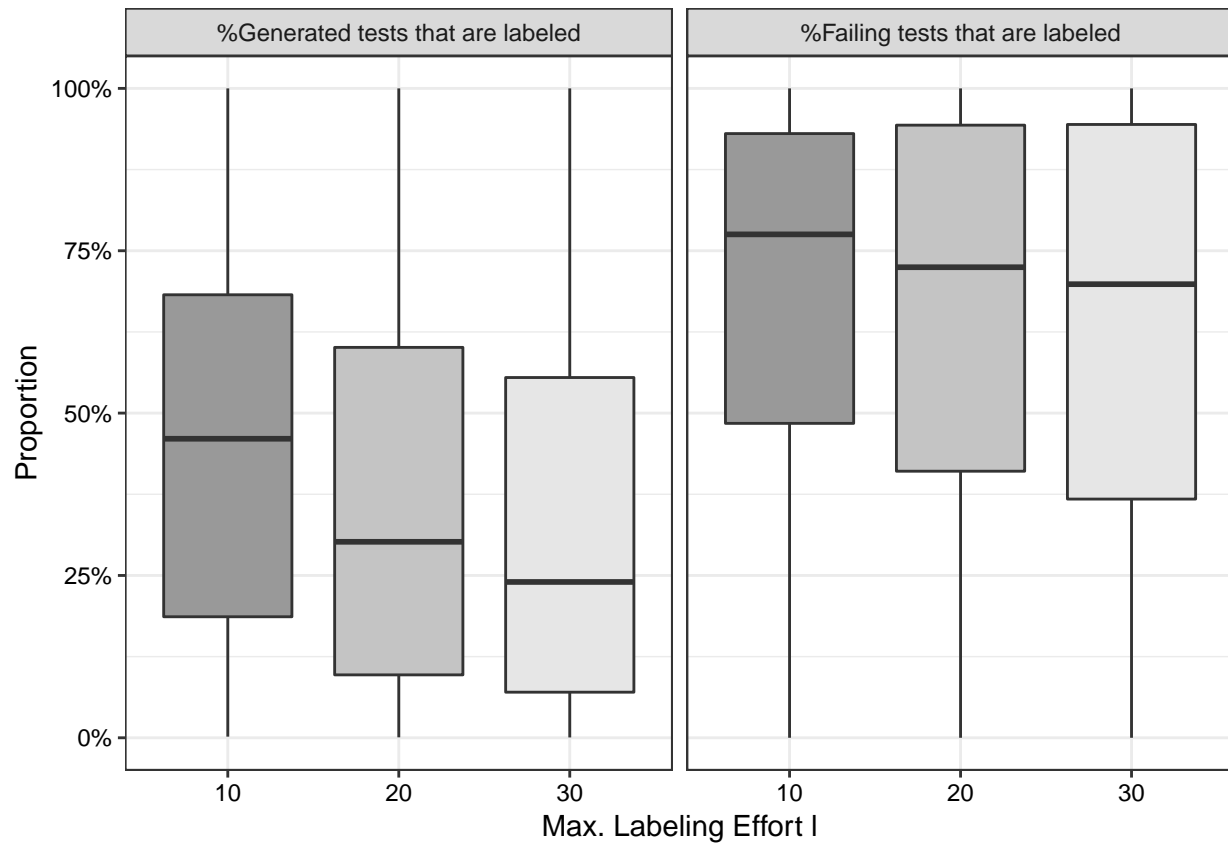0%

%Failing

```
## Saving 1.76 x 3.6 in image
```

## Human Effort

```
## [1] "Average/Median %Generated tests that are labeled (-l 10):  0.446983376546824 0.46047609623703"
## [1] "Average/Median %Generated tests that are labeled (-l 20):  0.364415420893335 0.301810865191147"
## [1] "Average/Median %Generated tests that are labeled (-l 30):  0.334512881418475 0.239959364919063"
## [1] "Average/Median %Failing tests that are labeled (-l 10):  0.67289962491367 0.77521186440678"
## [1] "Average/Median %Failing tests that are labeled (-l 20):  0.649864169641316 0.72463768115942"
## [1] "Average/Median %Failing tests that are labeled (-l 30):  0.634223708282012 0.698513111651798"
## [1] "Average/Median Improved (-l 10):  4.73989209702213 1.43816233936533"
## [1] "Average/Median Improved (-l 20):  9.40043022903786 1.77113543091655"
## [1] "Average/Median Improved (-l 30):  11.0075321600529 1.94688840121631"
## [1] "Average/Median Prob. to generate a failing test (-l 10):  0.302922018638272 0.220238124253577"
## [1] "Average/Median Prob. to generate a failing test (-l 20):  0.302967165125857 0.223314606741573"
## [1] "Average/Median Prob. to generate a failing test (-l 30):  0.301374546763708 0.209925373134328"
## [1] "Average/Median Prob. to label a failing test (-l 10):  0.414021378335148 0.363333333333333"
## [1] "Average/Median Prob. to label a failing test (-l 20):  0.488738809278124 0.486910994764398"
## [1] "Average/Median Prob. to label a failing test (-l 30):  0.523769581085046 0.55272774852991"
```
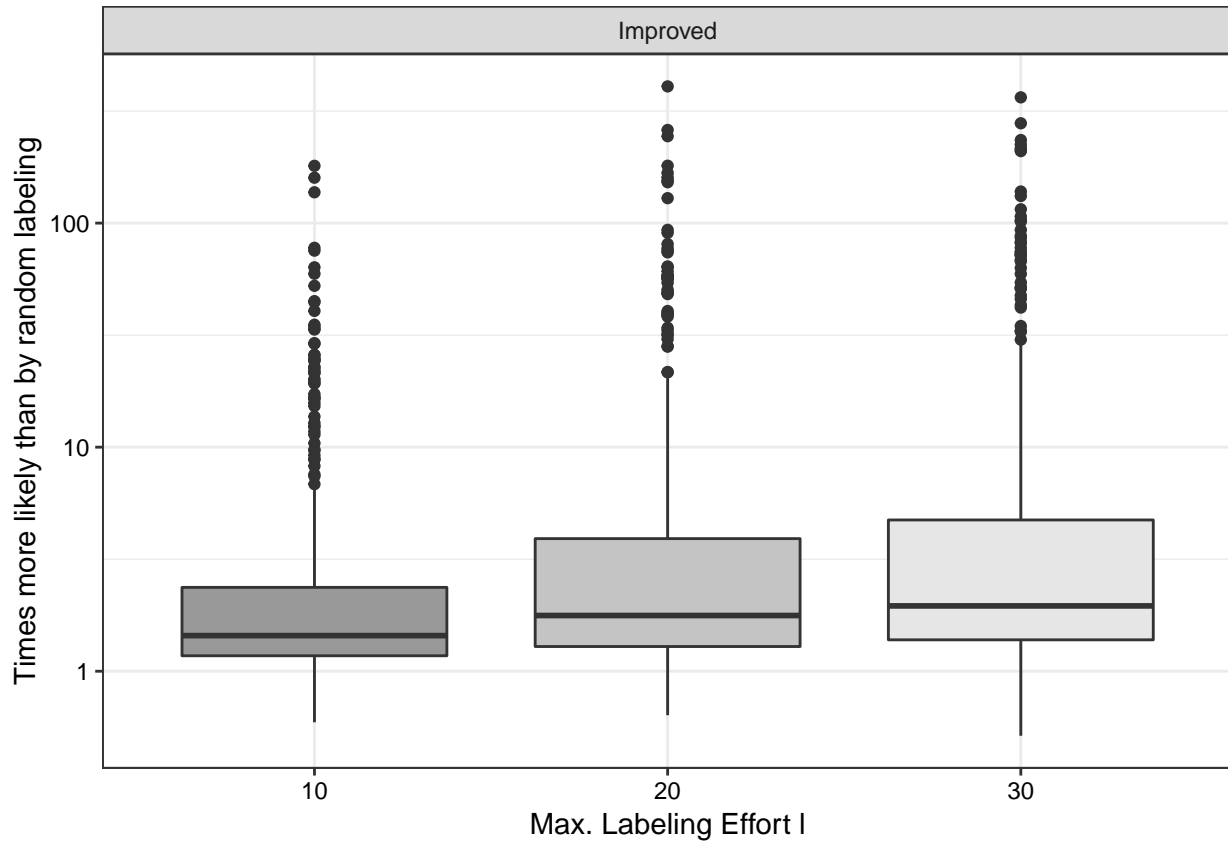
## Saving 5.2 x 3.6 in image

## Warning: Removed 12 rows containing non-finite values (stat_boxplot).

```
## Saving 5.2 x 3.6 in image

## Warning: Removed 12 rows containing non-finite values (stat_boxplot).

## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Removed 22 rows containing non-finite values (stat_boxplot).
```

## Patch Quality (Repairability)

```
## [1] "-l 10"

##                              V1          V2              V3
## 1-A-bug-18353198-18353306  :   30  Min.   : 1.0   Min.   :    9.0
## 104-A-bug-11691772-11691810:   30  1st Qu.: 8.0   1st Qu.:   13.0
## 104-A-bug-11691779-11691810:   30  Median :16.0   Median :   19.0
## 104-A-bug-11691794-11691810:   30  Mean   :15.5   Mean   :  213.5
## 104-A-bug-11939068-11939085:   30  3rd Qu.:23.0   3rd Qu.:   40.0
## 104-A-bug-13139540-13139570:   30  Max.   :30.0   Max.   : 8782.0
## (Other)                    :16110
##       V4              V5              V6              V7
## Min.   : 2.000  Min.   :    0.00  Min.   : 0.000  Min.   :   6.00
## 1st Qu.:10.000  1st Qu.:    2.00  1st Qu.: 1.000  1st Qu.: 35.00
## Median :10.000  Median :    6.00  Median : 4.000  Median : 46.00
## Mean   : 9.862  Mean   :   11.51  Mean   : 4.115  Mean   : 54.83
## 3rd Qu.:10.000  3rd Qu.:    9.00  3rd Qu.: 7.000  3rd Qu.: 63.00
## Max.   :10.000  Max.   : 3231.00  Max.   :10.000  Max.   :293.00
##
##       V8              V9              V10             V11
## Min.   :  0.00  Min.   :  1.00  Min.   :  0.000  manual:16290
## 1st Qu.: 23.00  1st Qu.:  4.00  1st Qu.:  2.000
## Median : 34.00  Median :  8.00  Median :  4.000
## Mean   : 41.13  Mean   : 13.08  Mean   :  9.407
```

```
## 3rd Qu.: 47.00   3rd Qu.: 17.00   3rd Qu.: 11.000
## Max.   :292.00   Max.   :101.00   Max.   :101.000
##
##                  V12              V13              V14
## ????              :1286   Min.   :  0.066   Min.   :  5.00
## NOREPAIR          :8804   1st Qu.:  3.469   1st Qu.: 26.00
## REPAIR            :4091   Median : 22.971   Median : 35.00
## TIMEOUT           :1900   Mean   :141.906   Mean   : 41.92
## VERIFICATIONFAILED: 209   3rd Qu.:209.204   3rd Qu.: 43.00
##                           Max.   :600.000   Max.   :178.00
##                                             NA's   :12199
##        V15              V16              V17
## Min.   : 17.00   Min.   : 17.0   autogen:16290
## 1st Qu.: 34.00   1st Qu.: 34.0
## Median : 42.00   Median : 42.0
## Mean   : 49.09   Mean   : 49.3
## 3rd Qu.: 51.00   3rd Qu.: 51.0
## Max.   :178.00   Max.   :178.0
## NA's   :12199    NA's   :12199
##                  V18              V19              V20
## ????              : 692   Min.   :  0.252   Min.   :  0.00
## NOREPAIR          :9826   1st Qu.: 28.464   1st Qu.: 25.00
## REPAIR            :3277   Median : 85.097   Median : 35.00
## TIMEOUT           :1418   Mean   :161.381   Mean   : 39.95
## VERIFICATIONFAILED:1077   3rd Qu.:217.908   3rd Qu.: 45.00
##                           Max.   :600.000   Max.   :178.00
##                                             NA's   :13013
##        V21              V22
## Min.   :  0.00   Min.   : 10.00
## 1st Qu.: 34.00   1st Qu.: 34.00
## Median : 42.00   Median : 42.00
## Mean   : 45.67   Mean   : 46.55
## 3rd Qu.: 50.00   3rd Qu.: 51.00
## Max.   :178.00   Max.   :178.00
## NA's   :13013    NA's   :13013

## [1] "-l 20"

##                                   V1            V2               V3
## 1-A-bug-18353198-18353306  :   30   Min.   : 1.00   Min.   :    18.0
## 104-A-bug-11691772-11691810:   30   1st Qu.: 7.00   1st Qu.:    32.0
## 104-A-bug-11691779-11691810:   30   Median :15.00   Median :    58.0
## 104-A-bug-11691794-11691810:   30   Mean   :14.98   Mean   :   546.7
## 104-A-bug-11939068-11939085:   30   3rd Qu.:23.00   3rd Qu.:   162.0
## 104-A-bug-13139540-13139570:   30   Max.   :30.00   Max.   : 19041.0
## (Other)                    :12952
##        V4              V5               V6               V7
## Min.   : 2.00   Min.   :   0.00   Min.   : 0.000   Min.   :  6.00
## 1st Qu.:20.00   1st Qu.:   8.00   1st Qu.: 4.000   1st Qu.: 35.00
## Median :20.00   Median :  15.00   Median :10.000   Median : 46.00
## Mean   :19.14   Mean   :  24.87   Mean   : 9.549   Mean   : 53.63
## 3rd Qu.:20.00   3rd Qu.:  19.00   3rd Qu.:15.000   3rd Qu.: 62.00
## Max.   :20.00   Max.   :3826.00   Max.   :20.000   Max.   :293.00
##
##        V8              V9              V10              V11
```

```
## Min.    :   2.0   Min.    :   1.00   Min.    :   0.000   manual:13132
## 1st Qu.: 26.0   1st Qu.:   4.00   1st Qu.:   2.000
## Median : 37.0   Median :   8.00   Median :   4.000
## Mean    : 43.7   Mean    : 12.94   Mean    :   9.389
## 3rd Qu.: 49.0   3rd Qu.: 17.00   3rd Qu.: 11.000
## Max.    :291.0   Max.    :101.00   Max.    :101.000
##
##    V12              V13              V14              V15
## Mode:logical    Mode:logical    Mode:logical    Mode:logical
## NA's:13132      NA's:13132      NA's:13132      NA's:13132
##
##
##
##
##
##    V16                 V17               V18              V19
## Mode:logical    autogen:13132    ????     : 745   Min.    :   0.329
## NA's:13132                       NOREPAIR:7741   1st Qu.: 77.397
##                                  REPAIR   :2162   Median :187.037
##                                  TIMEOUT :2484   Mean    :258.409
##                                                   3rd Qu.:449.139
##                                                   Max.    :600.000
##
##     V20                V21                V22
## Min.    :   0.00   Min.    :   0.00   Min.    : 15.00
## 1st Qu.: 26.00   1st Qu.: 34.00   1st Qu.: 34.00
## Median : 40.00   Median : 42.00   Median : 43.00
## Mean    : 42.58   Mean    : 46.64   Mean    : 47.61
## 3rd Qu.: 51.00   3rd Qu.: 54.00   3rd Qu.: 54.00
## Max.    :178.00   Max.    :178.00   Max.    :178.00
## NA's  :10970   NA's  :10970   NA's  :10970
##
## [1] "-l 30"
##
##                                      V1               V2              V3
## 1-A-bug-18353198-18353306 :  30   Min.    :  1.00   Min.    :      19.0
## 104-A-bug-11691772-11691810:  30   1st Qu.:  5.00   1st Qu.:      50.0
## 104-A-bug-11691779-11691810:  30   Median :14.00   Median :     107.0
## 104-A-bug-11691794-11691810:  30   Mean    :14.17   Mean    :    803.4
## 104-A-bug-11939068-11939085:  30   3rd Qu.:22.00   3rd Qu.:   360.2
## 104-A-bug-13139540-13139570:  30   Max.    :30.00   Max.    :16090.0
## (Other)                      :9636
##       V4               V5               V6               V7
## Min.    : 2.00   Min.    :   0.00   Min.    :   0.00   Min.    :   6.00
## 1st Qu.:30.00   1st Qu.: 13.00   1st Qu.: 5.00   1st Qu.: 35.00
## Median :30.00   Median : 24.00   Median :16.00   Median : 46.00
## Mean    :27.46   Mean    : 34.87   Mean    :14.86   Mean    : 54.38
## 3rd Qu.:30.00   3rd Qu.: 29.00   3rd Qu.:25.00   3rd Qu.: 62.00
## Max.    :30.00   Max.    :3085.00   Max.    :30.00   Max.    :293.00
##
##       V8               V9               V10              V11
## Min.    :   2.00   Min.    :   1.00   Min.    :   0.000   manual:9816
## 1st Qu.: 28.00   1st Qu.:   3.00   1st Qu.:   2.000
## Median : 38.00   Median :   8.00   Median :   4.000
## Mean    : 45.86   Mean    : 13.14   Mean    :   9.848
```
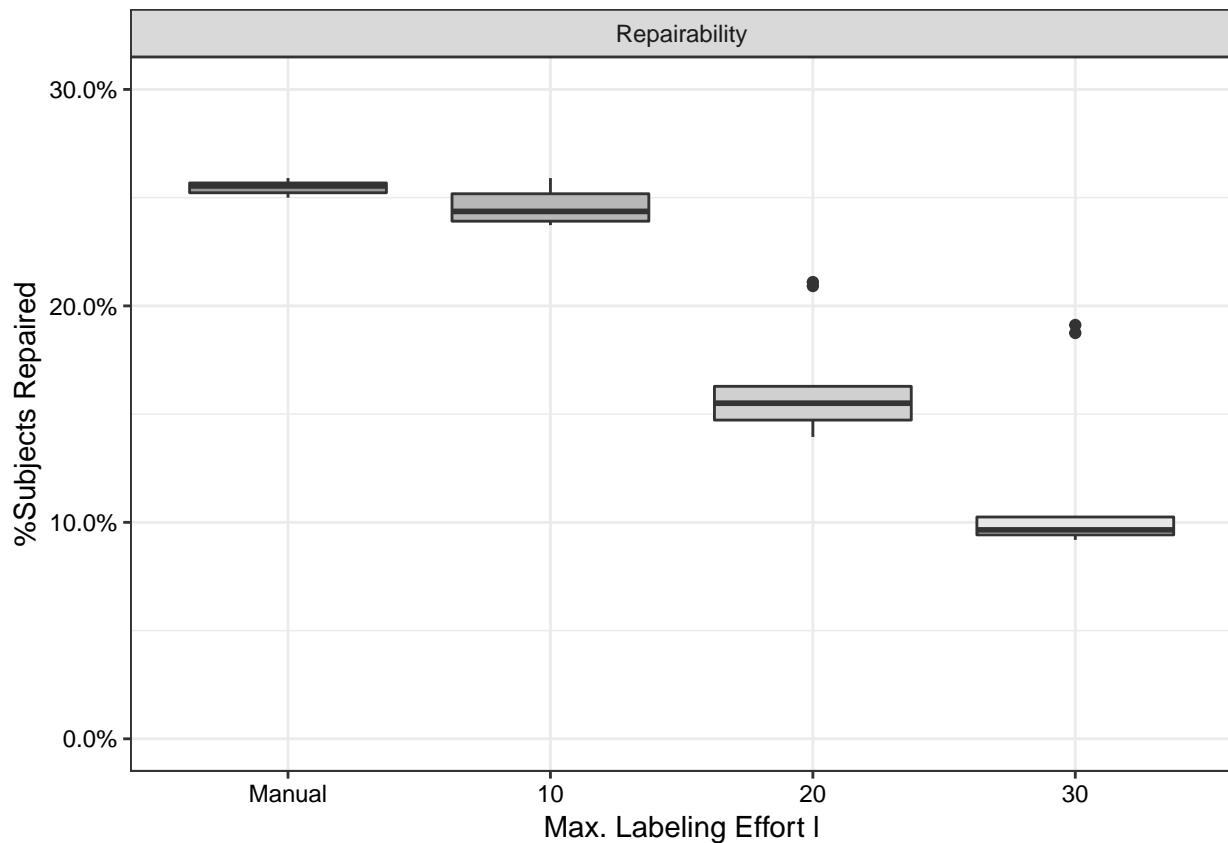
```
##  3rd Qu.: 50.00    3rd Qu.: 18.00    3rd Qu.: 12.000
##  Max.   :293.00    Max.    :101.00   Max.    :101.000
##
##      V12              V13               V14               V15
##  Mode:logical     Mode:logical      Mode:logical      Mode:logical
##  NA's:9816        NA's:9816         NA's:9816         NA's:9816
##
##
##
##
##
##      V16              V17               V18               V19
##  Mode:logical     autogen:9816      ????    : 473     Min.   :  0.29
##  NA's:9816                          NOREPAIR:5744     1st Qu.: 82.71
##                                     REPAIR  :1496     Median :184.92
##                                     TIMEOUT :2103     Mean   :270.42
##                                                       3rd Qu.:505.71
##                                                       Max.   :600.00
##
##      V20              V21               V22
##  Min.   :  0.00   Min.    :  0.00   Min.    : 15.00
##  1st Qu.: 26.00   1st Qu.: 34.75    1st Qu.: 35.00
##  Median : 40.00   Median : 42.00    Median : 43.00
##  Mean   : 42.11   Mean    : 46.32   Mean    : 47.41
##  3rd Qu.: 49.00   3rd Qu.: 53.00    3rd Qu.: 54.00
##  Max.   :178.00   Max.    :178.00   Max.    :178.00
##  NA's   :8320     NA's    :8320     NA's    :8320
```
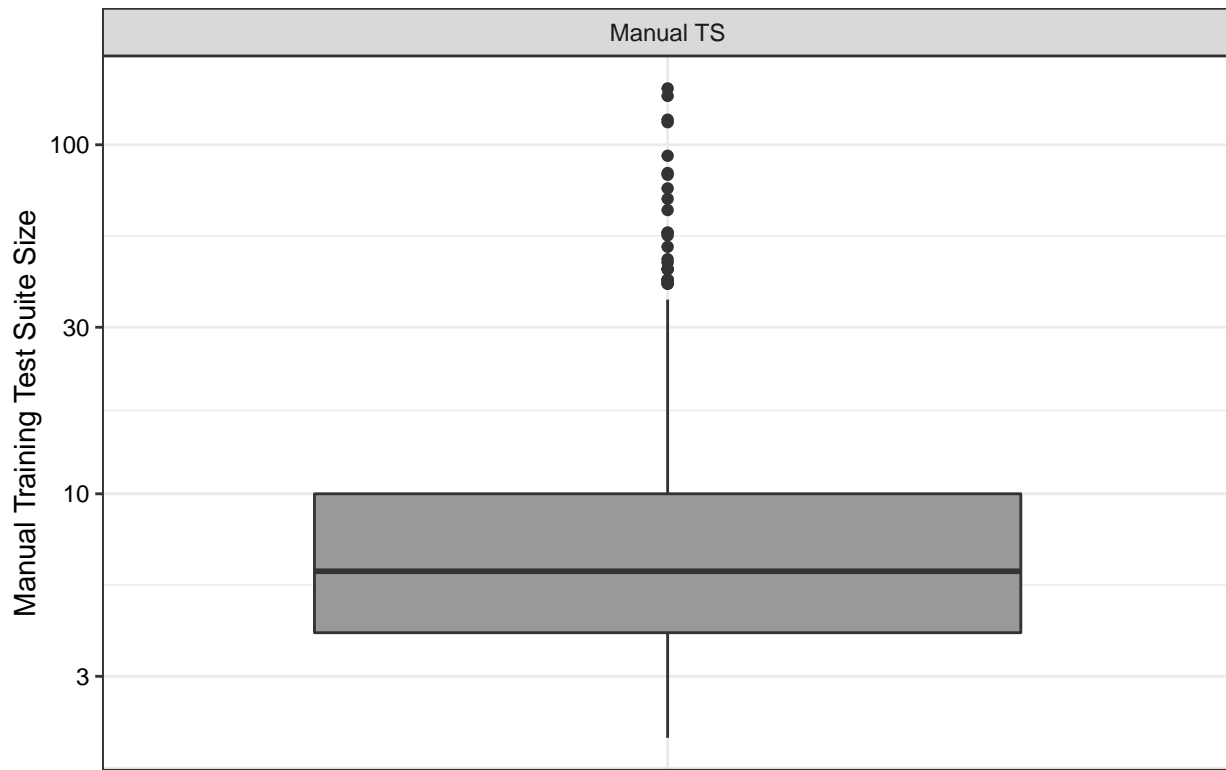
```
## Saving 4.55 x 3.15 in image
training = data.frame("subject"=character(), "max_label"=character(),
                      "variable"=character(),  "value" = numeric())
d = read.table("manualTrainingTSsize.csv",sep=",",comment.char = "#")
for (Subject in levels(factor(d$V1))){
  specific = subset(d, V1 == Subject)
  training <- rbind (training, data.frame(subject = Subject, max_label="",
                                          variable = "Manual TS",
                                          value = specific$V2))

}

print(paste("Average/Median training test suite: ",
            mean(training$value, na.rm=TRUE),
            median(training$value, na.rm=TRUE)))
```

```
## [1] "Average/Median training test suite:  10.7608695652174 6"
```

```
ggplot(training, aes(max_label, value)) +
  geom_boxplot(aes(fill=max_label)) +
  scale_y_log10() +
  facet_grid(~ variable) +
  xlab("") + ylab("Manual Training Test Suite Size") +
  theme(legend.position="none", legend.title= element_blank(),
        axis.text.x = element_text(colour = "black"), axis.text.y = element_text(colour = "black")) +
  scale_fill_grey(start = 0.6, end = .9)
```
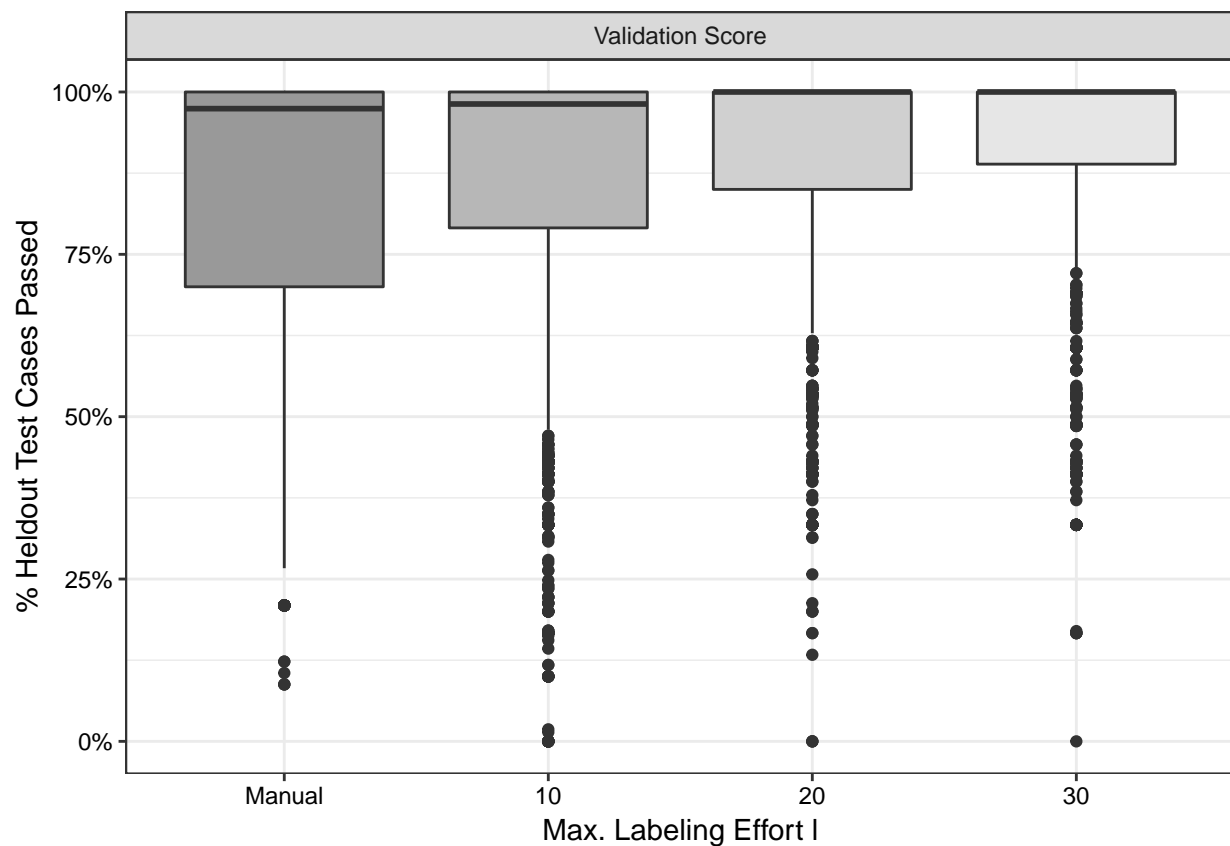


```
ggsave(filename = "training.pdf", width=2.2,scale=0.7)
```

```
## Saving 1.54 x 3.15 in image
```

# Patch Quality (Validation Test Suite: Together)

# Patch Quality (Validation Test Suite: Individually)

```
## [1] "l10 (manual): 4091"
```

```
## [1] "l10 (autogen): 3277"
```

```
## [1] "l20 (manual): 0"
```

```
## [1] "l20 (autogen): 2162"
```

```
## [1] "l30 (manual): 0"
```

```
## [1] "l30 (autogen): 1496"
```

```
## [1] "Average/Median Manual:  0.148537519535924 0.024390243902439"
```

```
## [1] "Average/Median -l 10:  0.138126389517177 0.015625"
```

```
## [1] "Average/Median -l 20:  0.101219289429628 0"
```

```
## [1] "Average/Median -l 30:  0.103172317158473 0"
```

```
## Warning: Removed 107 rows containing non-finite values (stat_boxplot).
```



```
## Saving 4.55 x 3.15 in image
```

```
## Warning: Removed 107 rows containing non-finite values (stat_boxplot).
```