Data analysis experiments - Information diffusion in complex emergencies

Jasper Meijering

26/02/2019

This document accompanies the thesis "Information diffusion in complex emergencies". This files does not aim to be self containing. Instead it aims to give the reader an idea of which analysis have been performed during the research project.

This analysis follows 5 steps. These steps are:

1) Data import and cleaning
2) Analysis of core assumptions
3) Analysis of individual strategies
4) Analysis of comprehensive strategies
5) Analayis for structural validation

The remainder of this document discusses the steps one by one.

1. Data import and cleaning

2. Analysis of core assumptions

Does the shocks assumption change the behaviour?

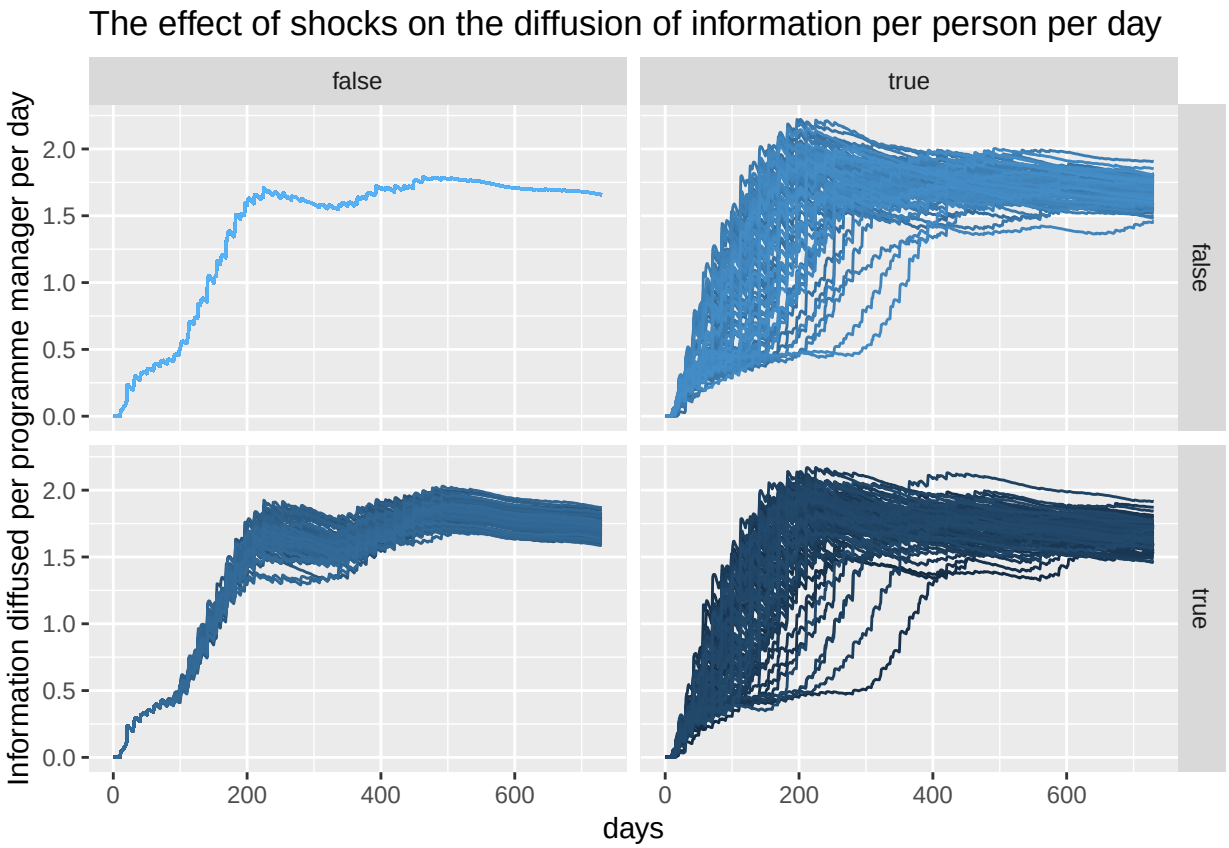
Figure 1 shows information diffusion for 50 identical disasters. The only variable that was changed for these disasters was the time at which need chancing shocks took place. This figure show that the level of information diffusion is different for various timings of shocks. This behaviour illustrates how the information landschap is continous evolving based on amongst others the timing of shocks.

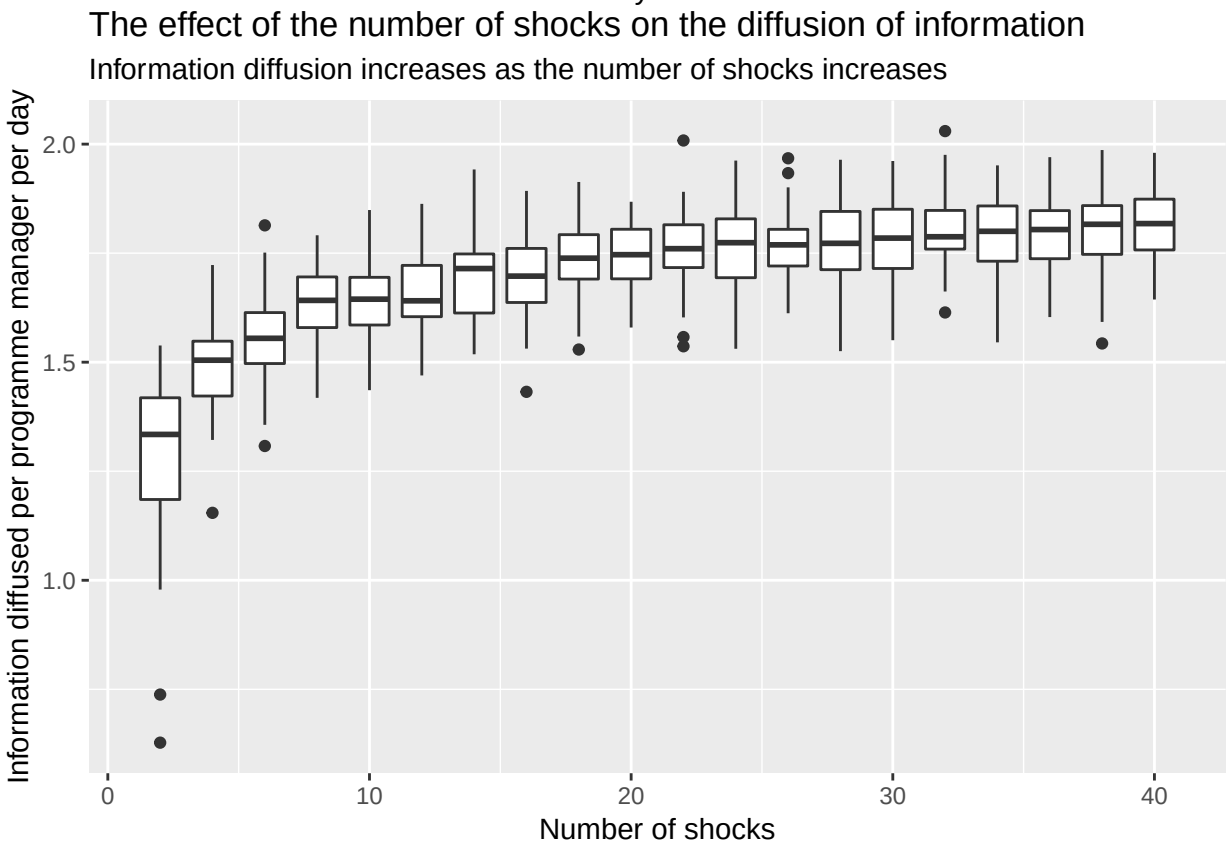
Figure 2 shows that the number of days worked increase as the number of shocks increase.

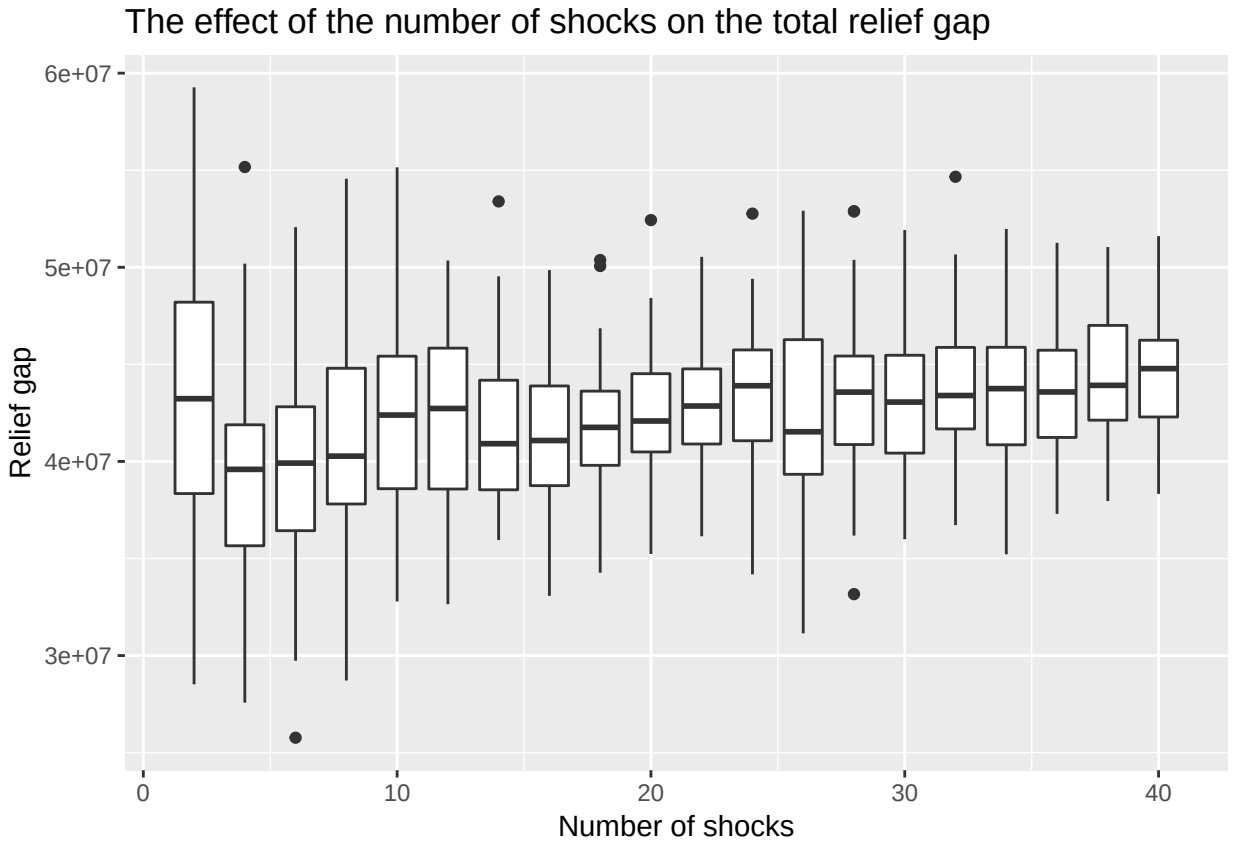
Figure 3 shows the effect of more shocks on the relief gap.

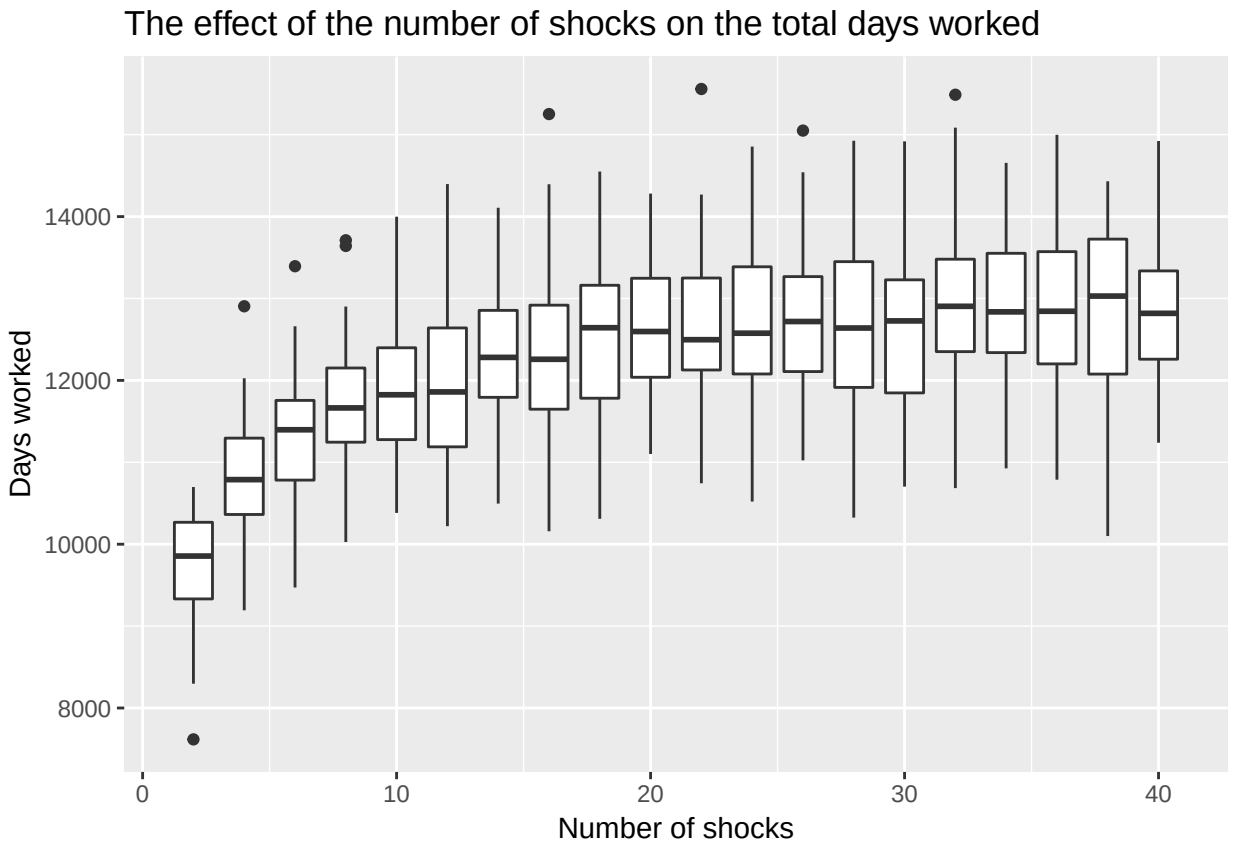
Figure 4 shows the effect of the number of shocks on the diffusion of information.

Figure 5 also shows the effect of the number of shocks but then on the diffusion of information per programme manager per day.


The effect of shocks on the diffusion of information per person per day
false
true
false
true
Information diffused per programme manager per day
2.0
1.5
1.0
0.5
0.0
0
200
400
600
days

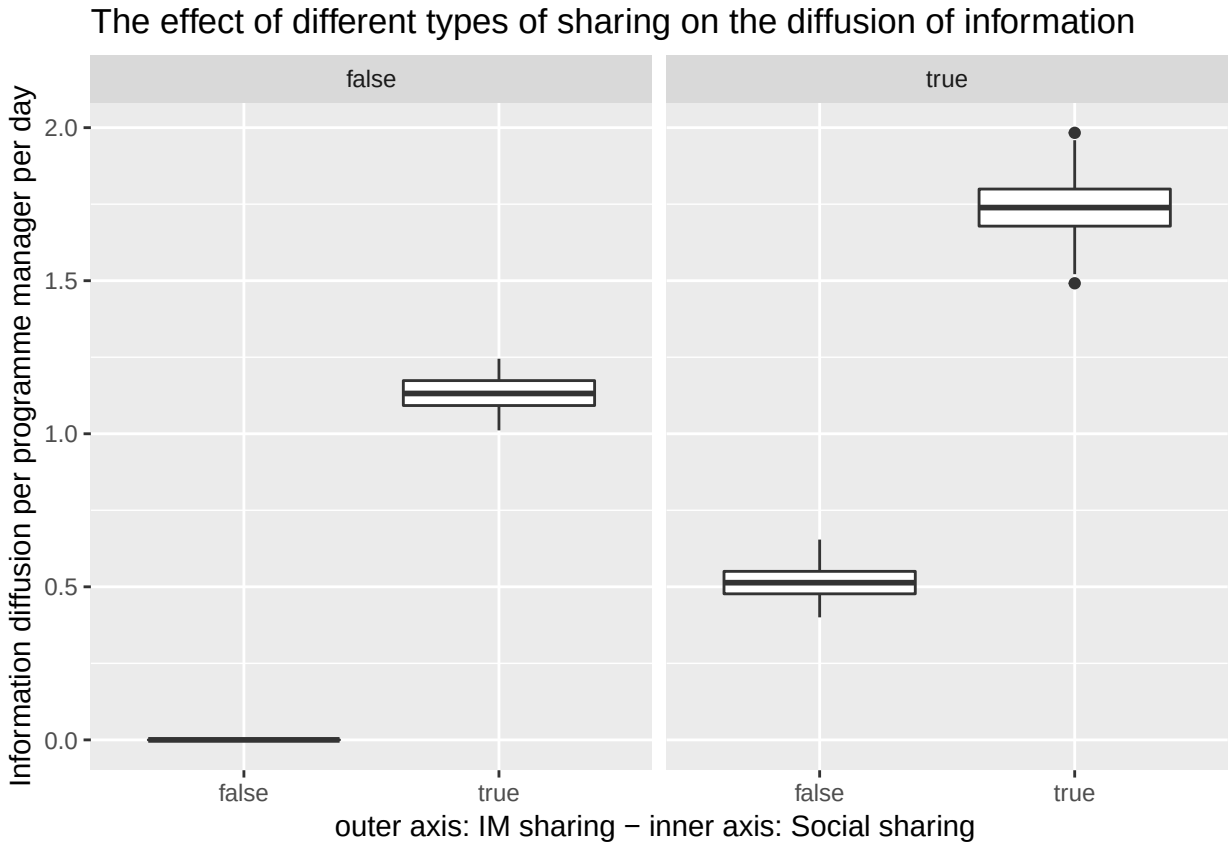

The effect of the number of shocks on the diffusion of information
Information diffusion increases as the number of shocks increases
Information diffused per programme manager per day
2.0
1.5
1.0
0
10
20
30
40
Number of shocks

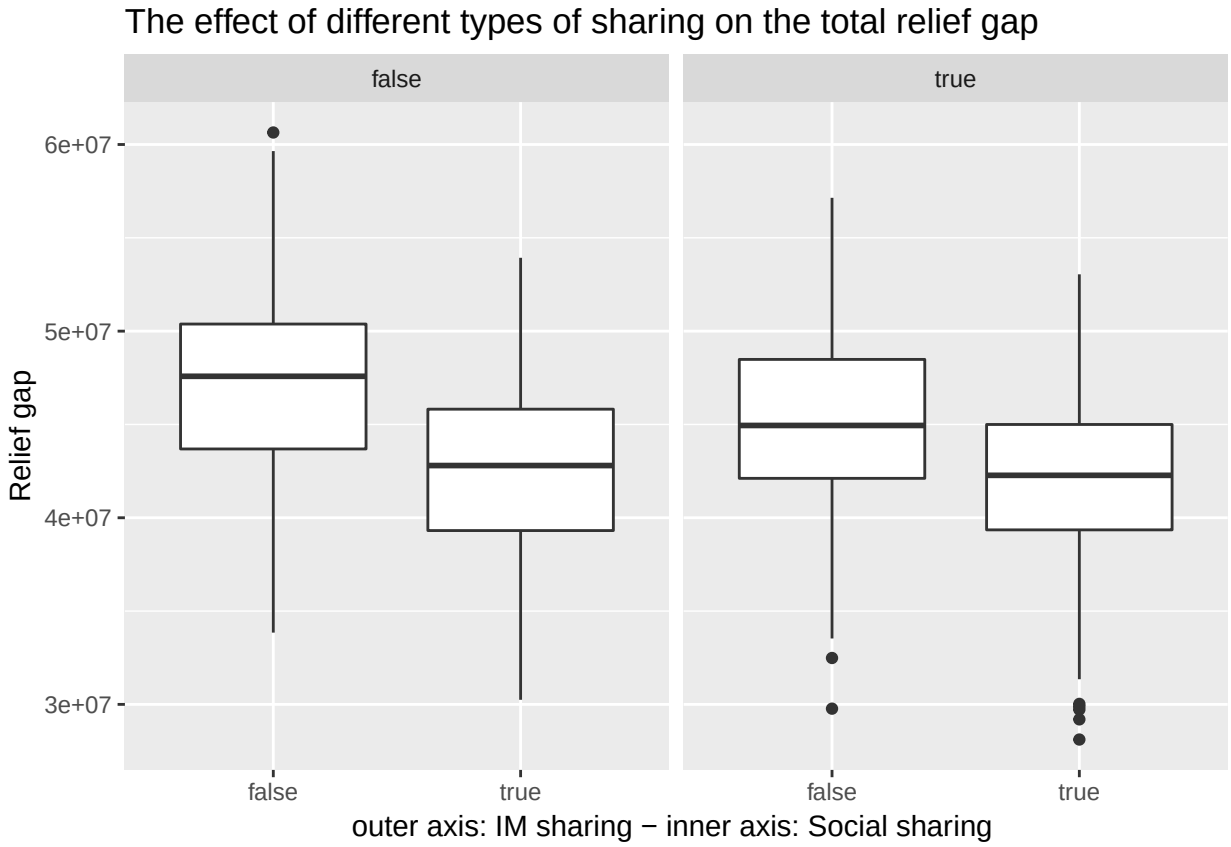

The effect of the number of shocks on the total relief gap
Relief gap
6e+07
5e+07
4e+07
3e+07
0
10
20
30
40
Number of shocks

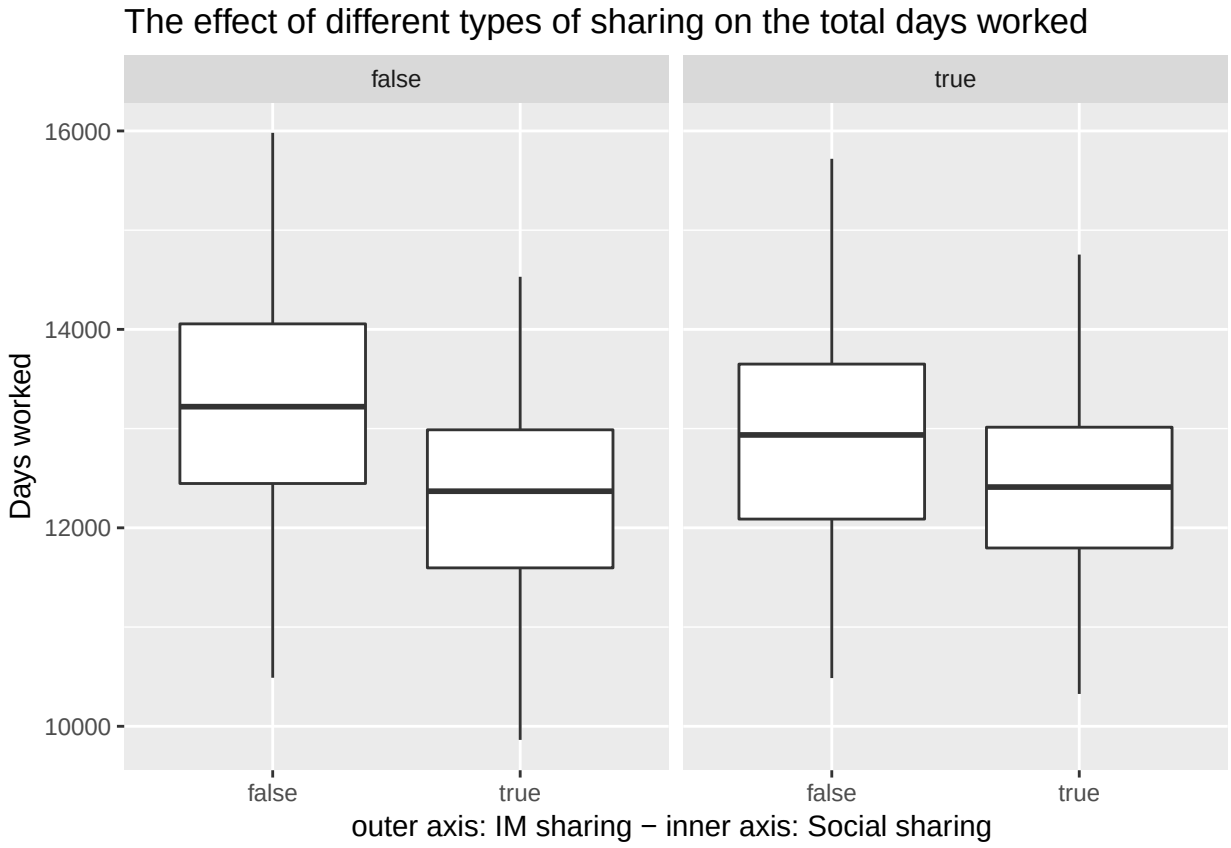

The effect of the number of shocks on the total days worked
Days worked
14000
12000
10000
8000
0
10
20
30
40
Number of shocks

```
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
```

Does the social network assumption change the behaviour?


The effect of different types of sharing on the diffusion of information
false
true
Information diffusion per programme manager per day
2.0
1.5
1.0
0.5
0.0
false
true
false
true
outer axis: IM sharing − inner axis: Social sharing


The effect of different types of sharing on the total relief gap
false
true
Relief gap
6e+07
5e+07
4e+07
3e+07
false
true
false
true
outer axis: IM sharing − inner axis: Social sharing



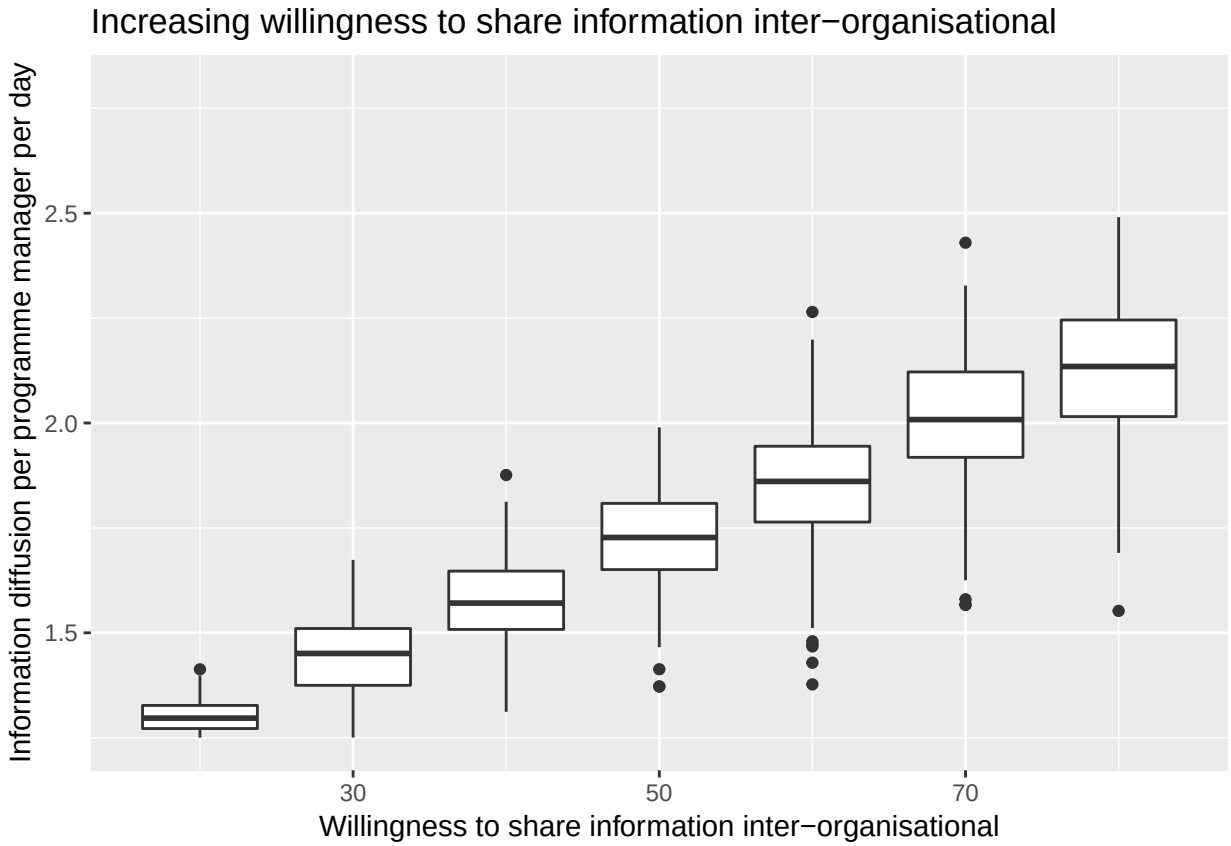

```
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
```

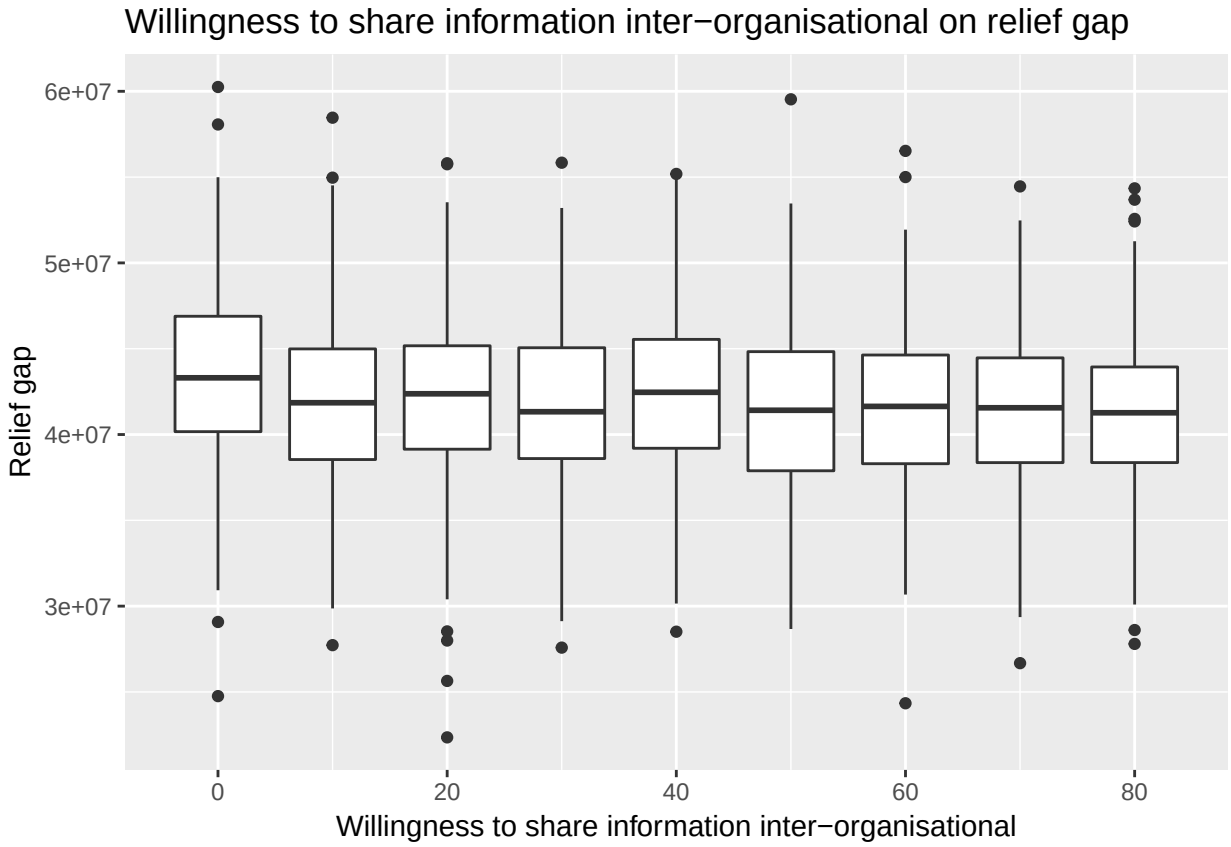
3. Analysis of individual strategies

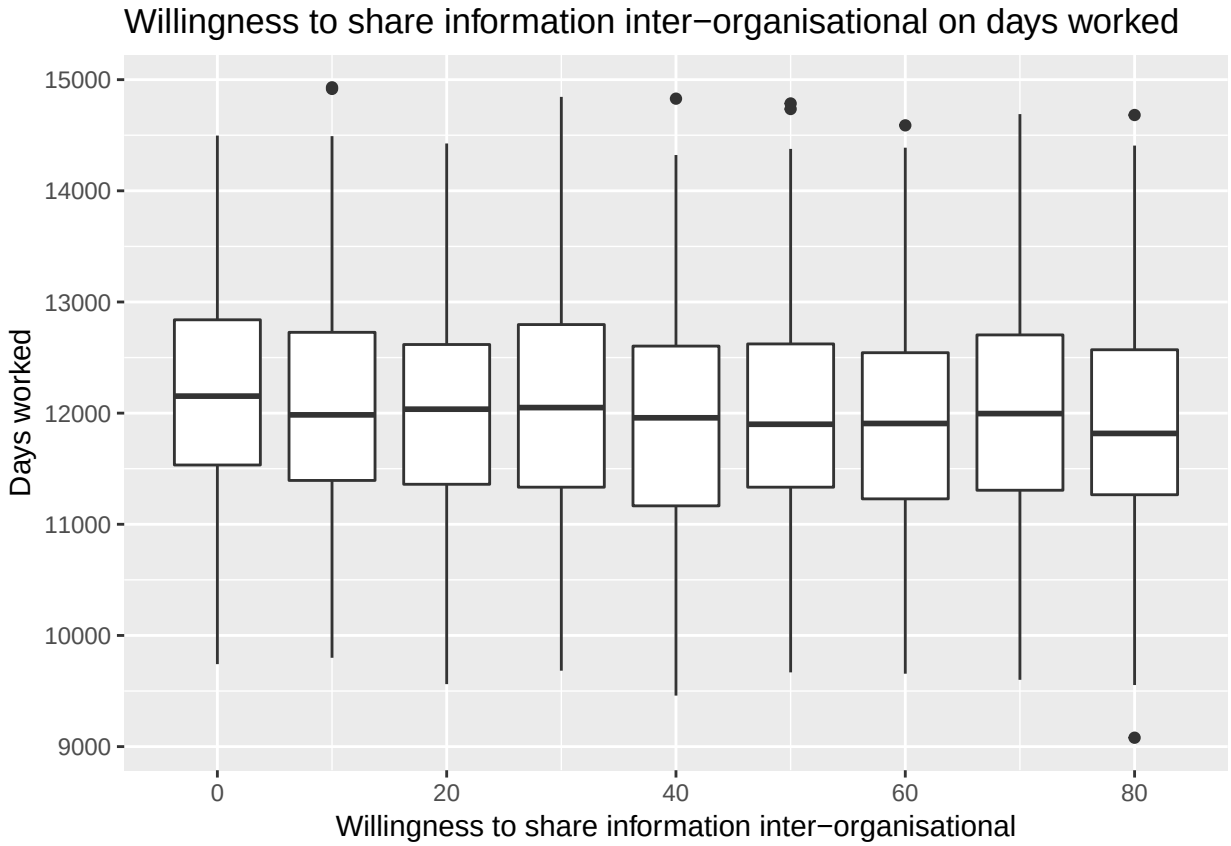
Willingness to share information inter-organisational

As willigningess to share information inter-organisational, the total information diffused increases. The effect on the other parameters is less clear. Also not that a willingness of 100% is inpossible especially in complex disasters."

```
## Warning: Removed 782 rows containing non-finite values (stat_boxplot).
```


Increasing willingness to share information inter−organisational
Information diffusion per programme manager per day
2.5
2.0
1.5
30
50
70
Willingness to share information inter−organisational


Willingness to share information inter−organisational on relief gap
Relief gap
6e+07
5e+07
4e+07
3e+07
0
20
40
60
80
Willingness to share information inter−organisational

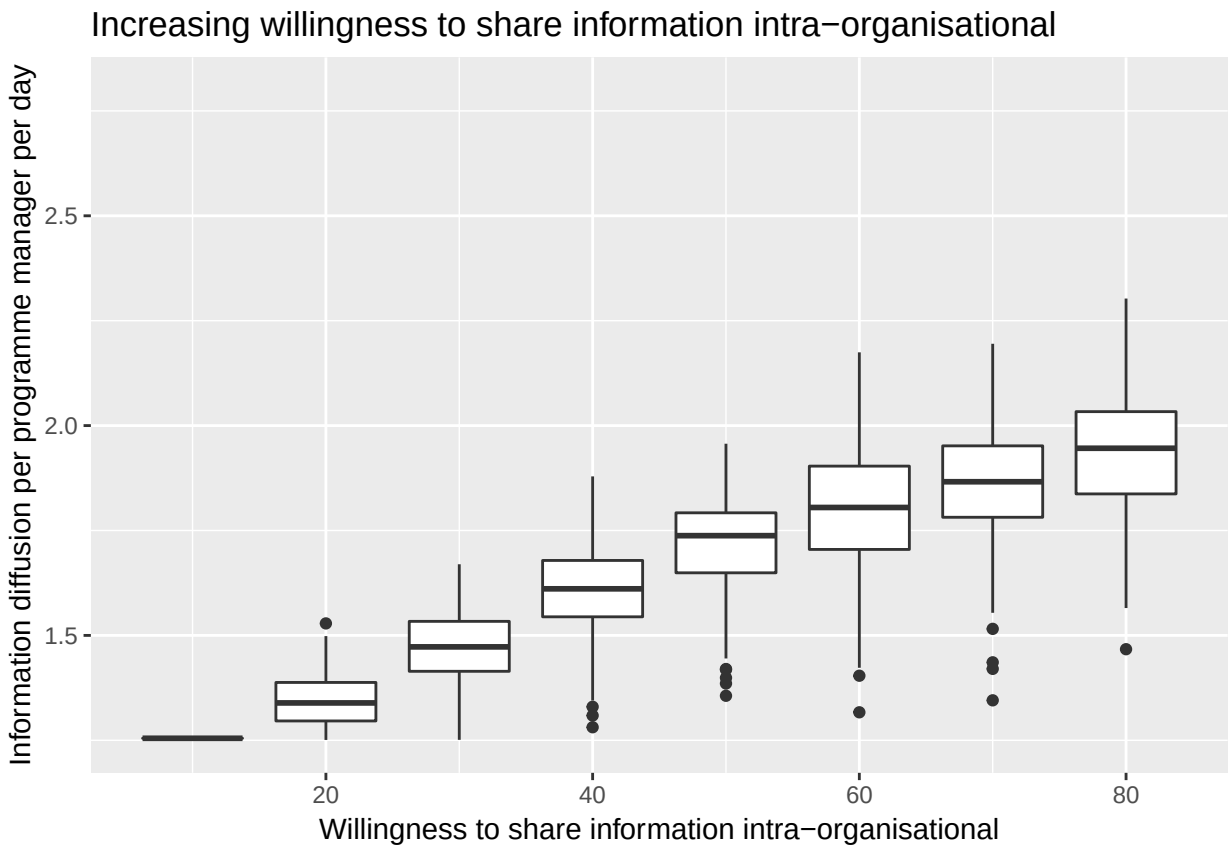


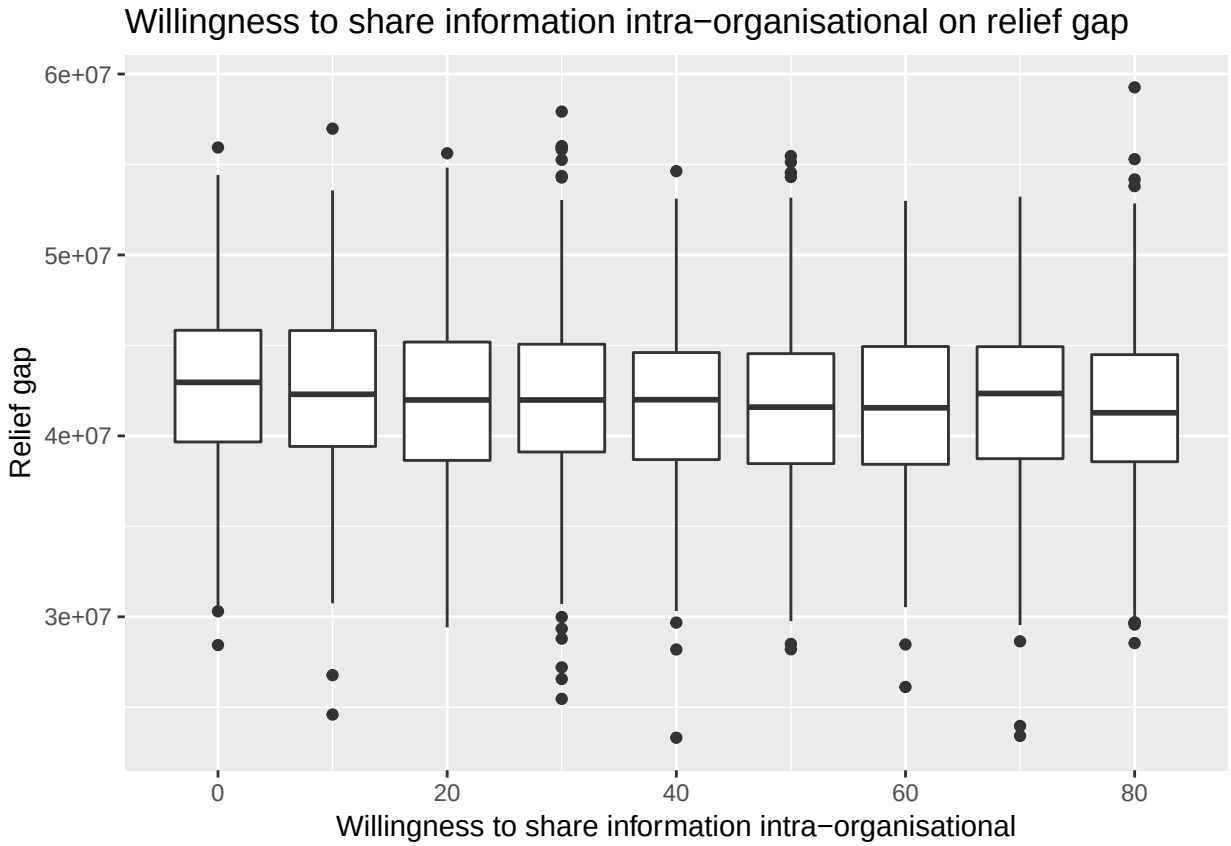

```
## Saving 6.5 x 4.5 in image
## Warning: Removed 782 rows containing non-finite values (stat_boxplot).
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
```

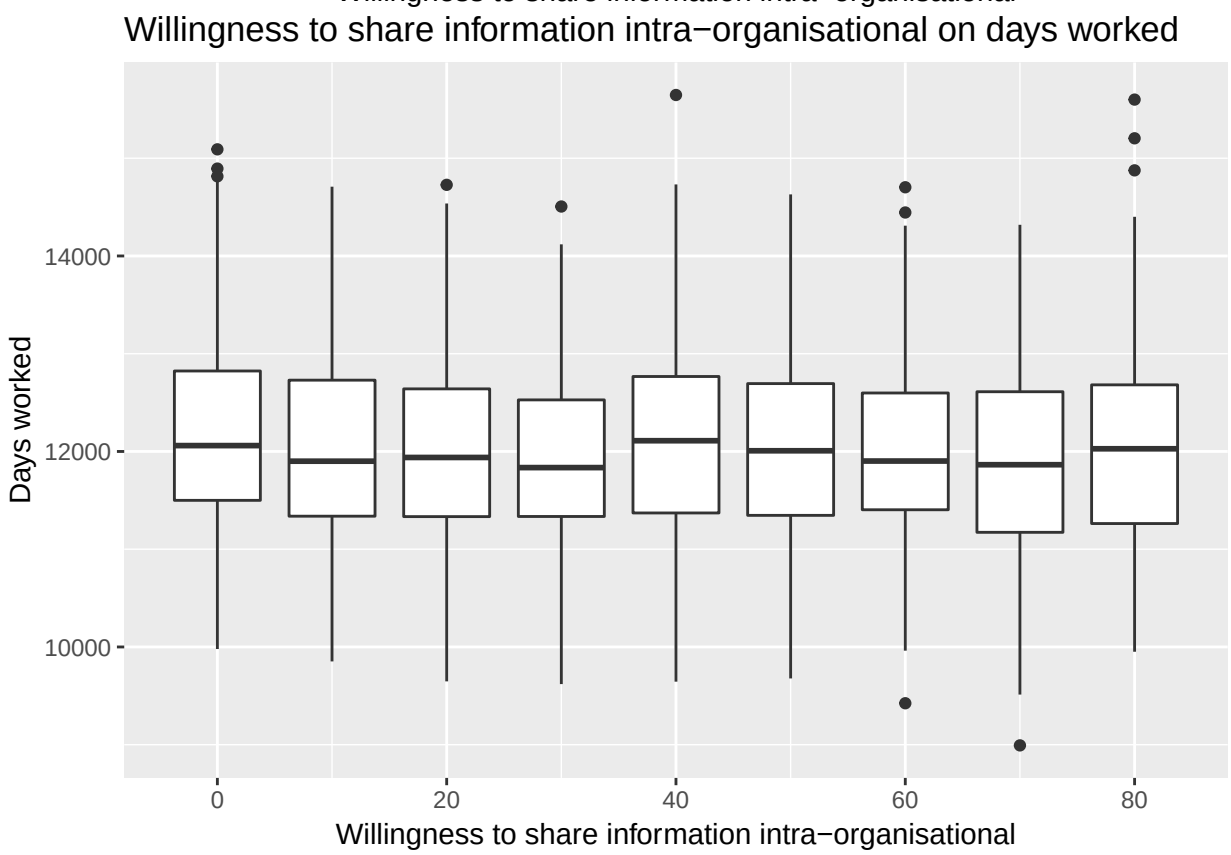
Willingness to share information intra-organisational

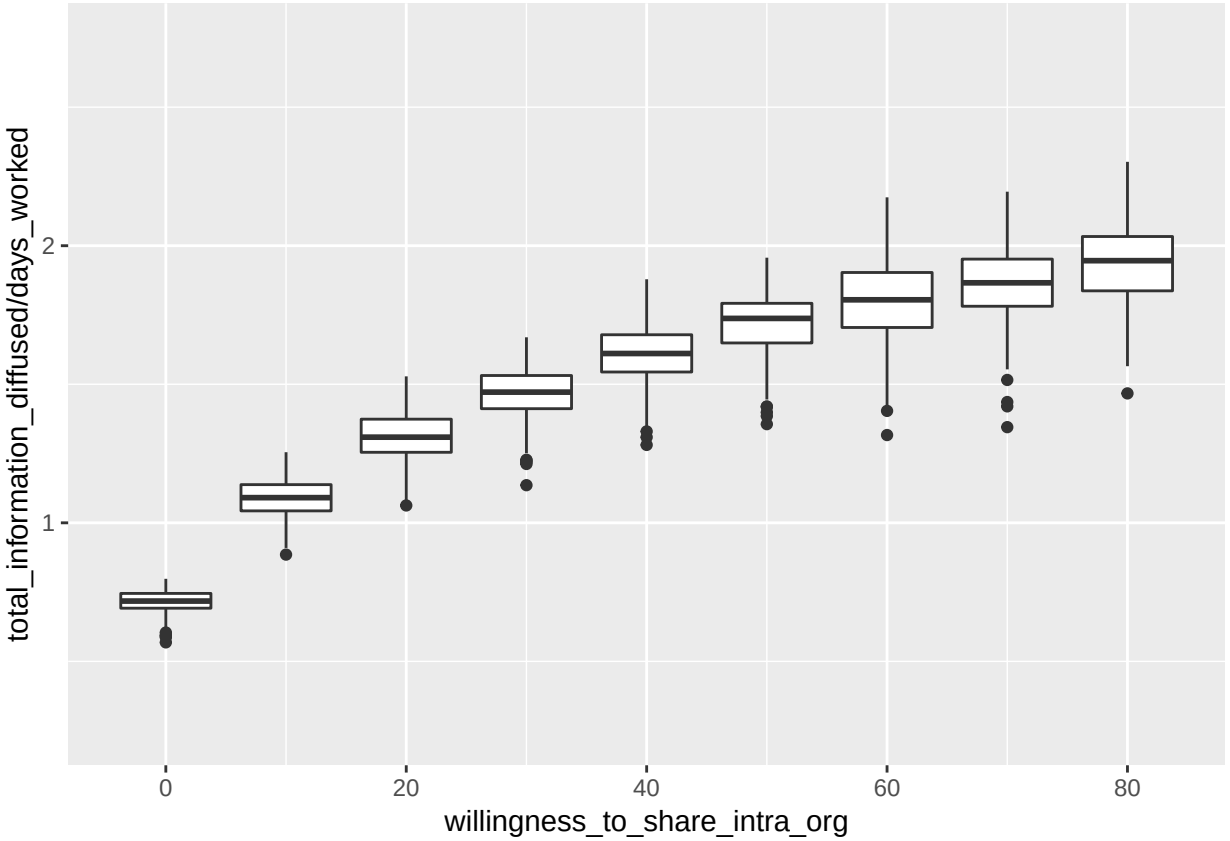
As willingingess to share information intra-organisational, the total information diffused increases but less as is the case for inter-organisational information sharing. The effect on the other parameters is less clear.

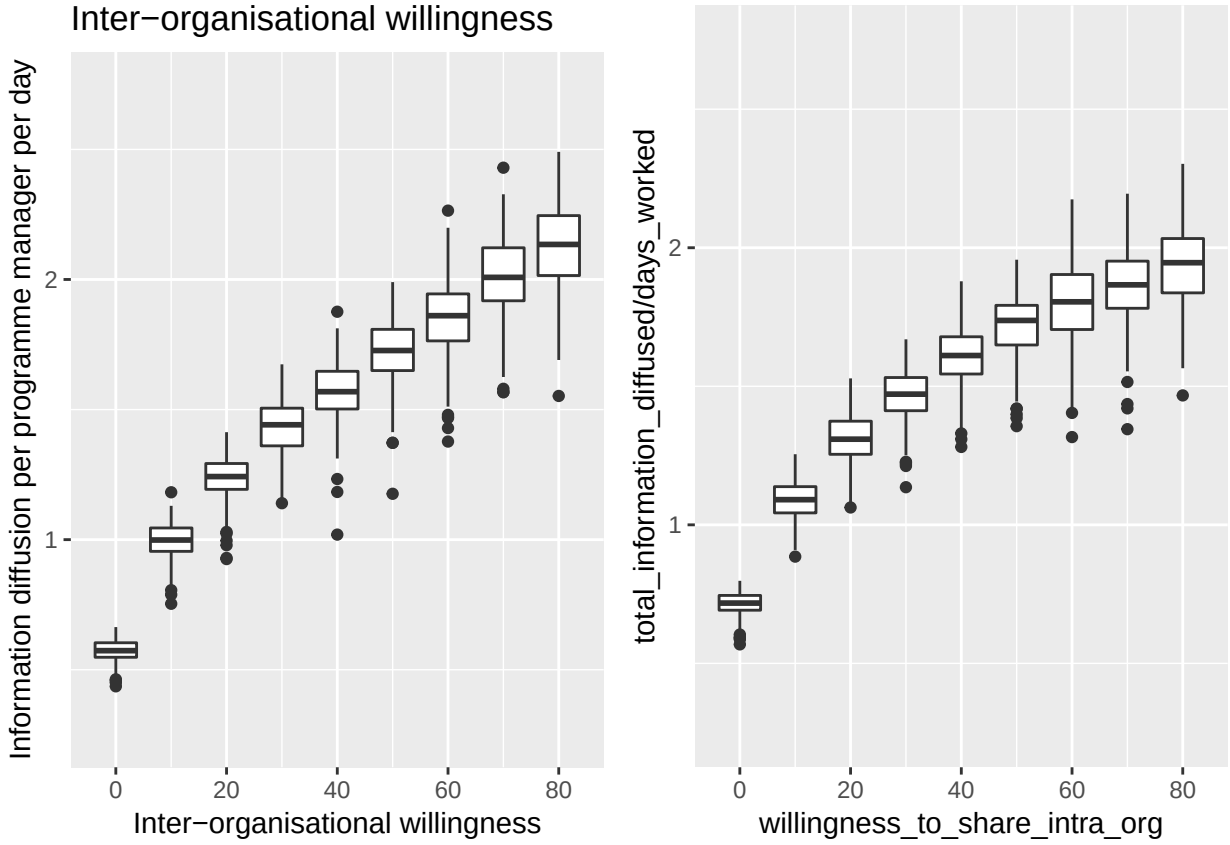
```
## Warning: Removed 673 rows containing non-finite values (stat_boxplot).
```




Willingness to share information intra−organisational on relief gap
Relief gap
6e+07
5e+07
4e+07
3e+07
0
20
40
60
80
Willingness to share information intra−organisational


Willingness to share information intra−organisational on days worked
Days worked
14000
12000
10000
0
20
40
60
80
Willingness to share information intra−organisational

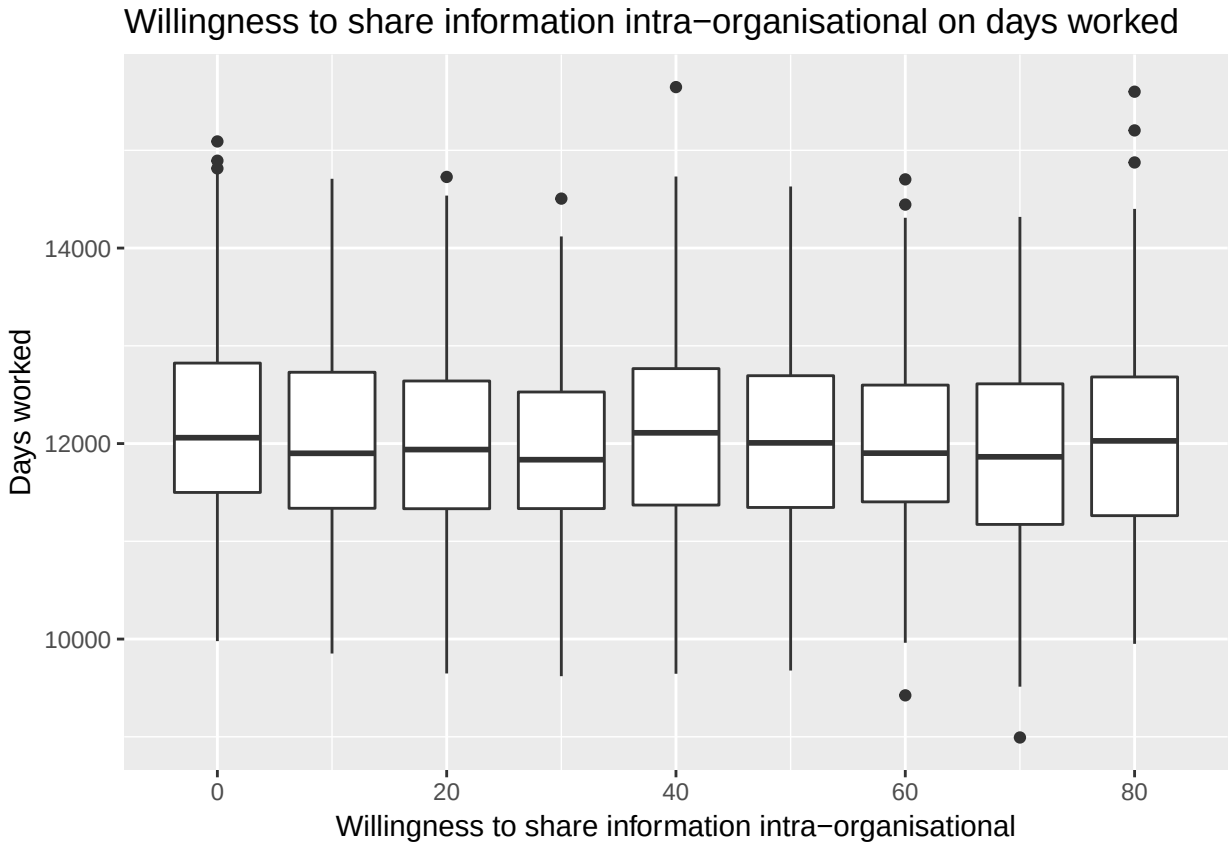





```
## Saving 6.5 x 4.5 in image
```

```
## Warning: Removed 673 rows containing non-finite values (stat_boxplot).
```

```
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
```

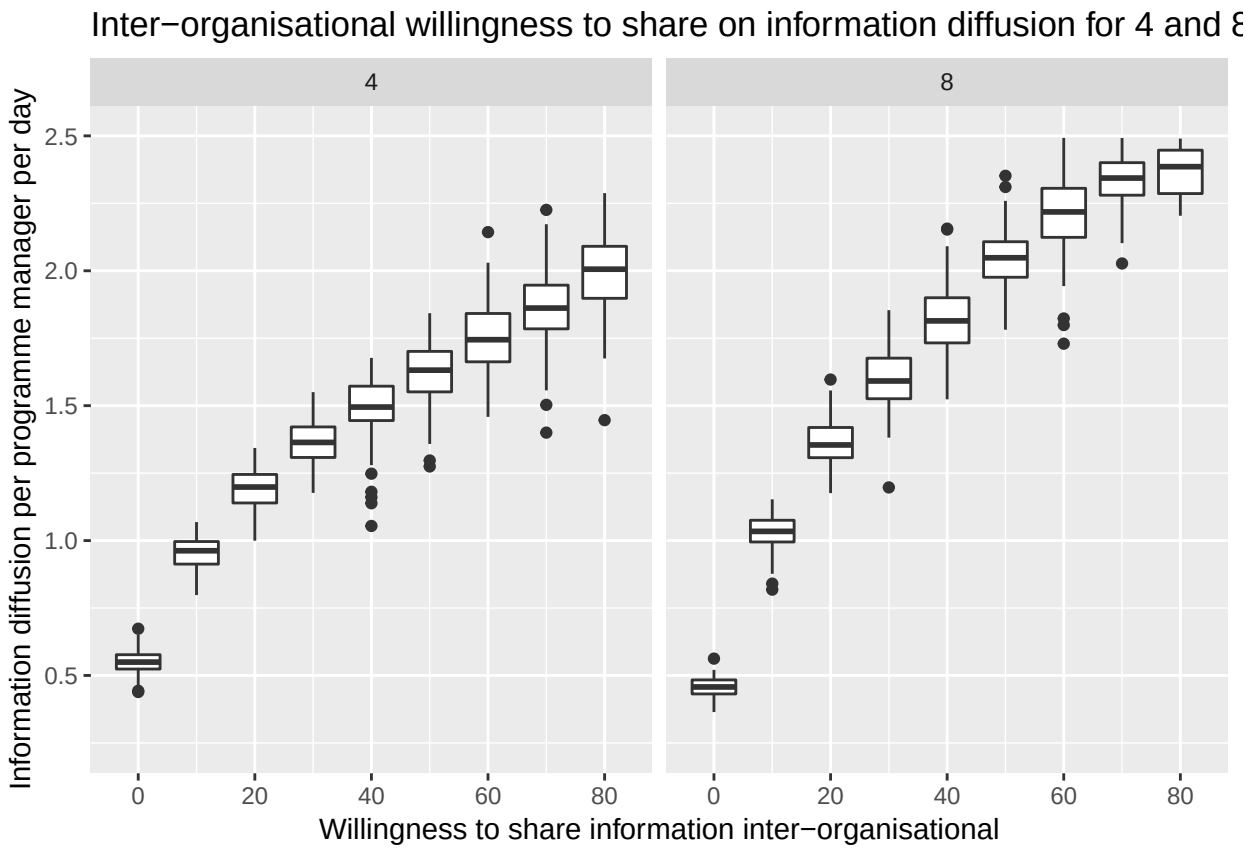


Sensitivty analysis willingness to share - Inter, intra and number of organisations.

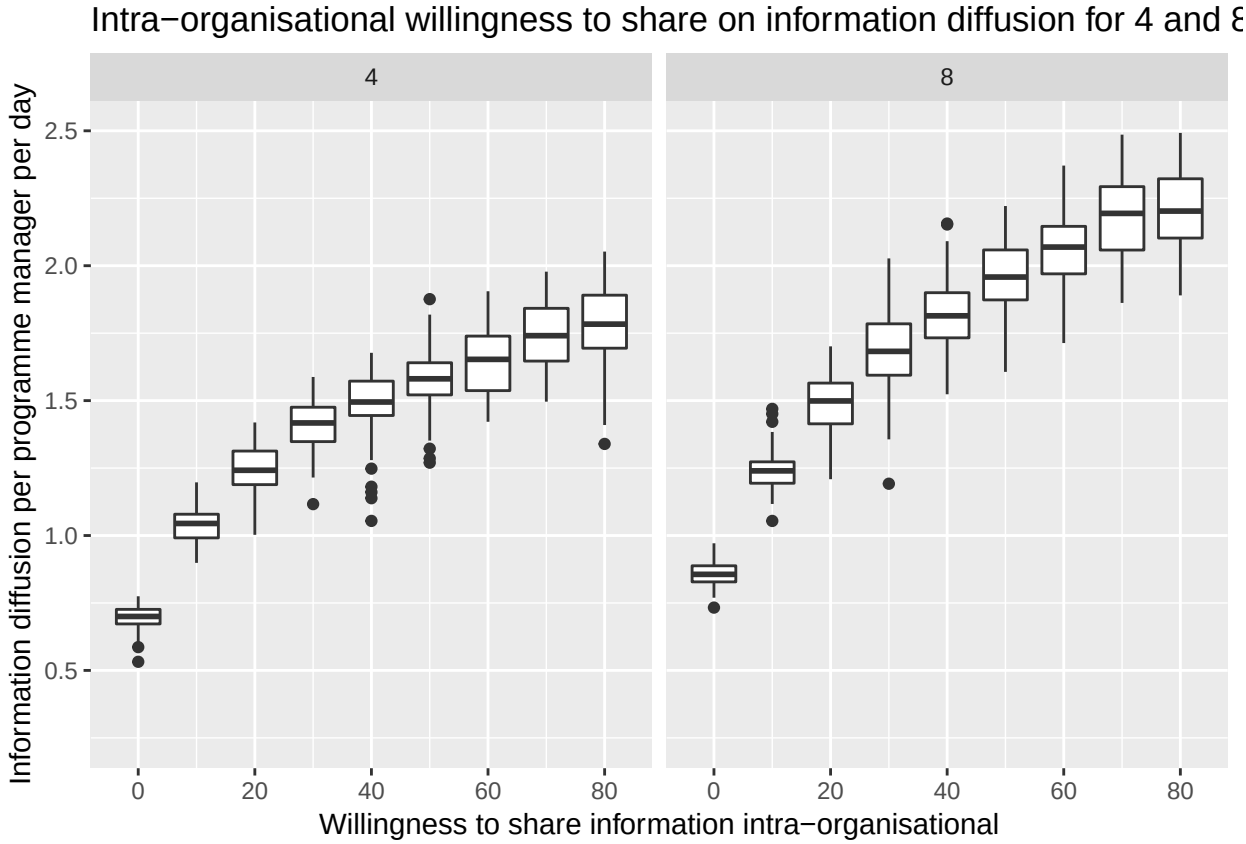
If number of organisations increase inf increases. If number of organsaitons increases the effect of the increasing inter organisations willingness to share information increases. Enforcess the conclusion that increasing inter-organisational willigness to share is more effective. It would be nice if I couls also evaluate the effect of having 12 organisations but that is computationally too intensive. -> future research.

Figures are not easy to interprete beceause of dubble effect. Effect of more organisations and effect of willingness to share.

```
## Warning: Removed 93 rows containing non-finite values (stat_boxplot).
```



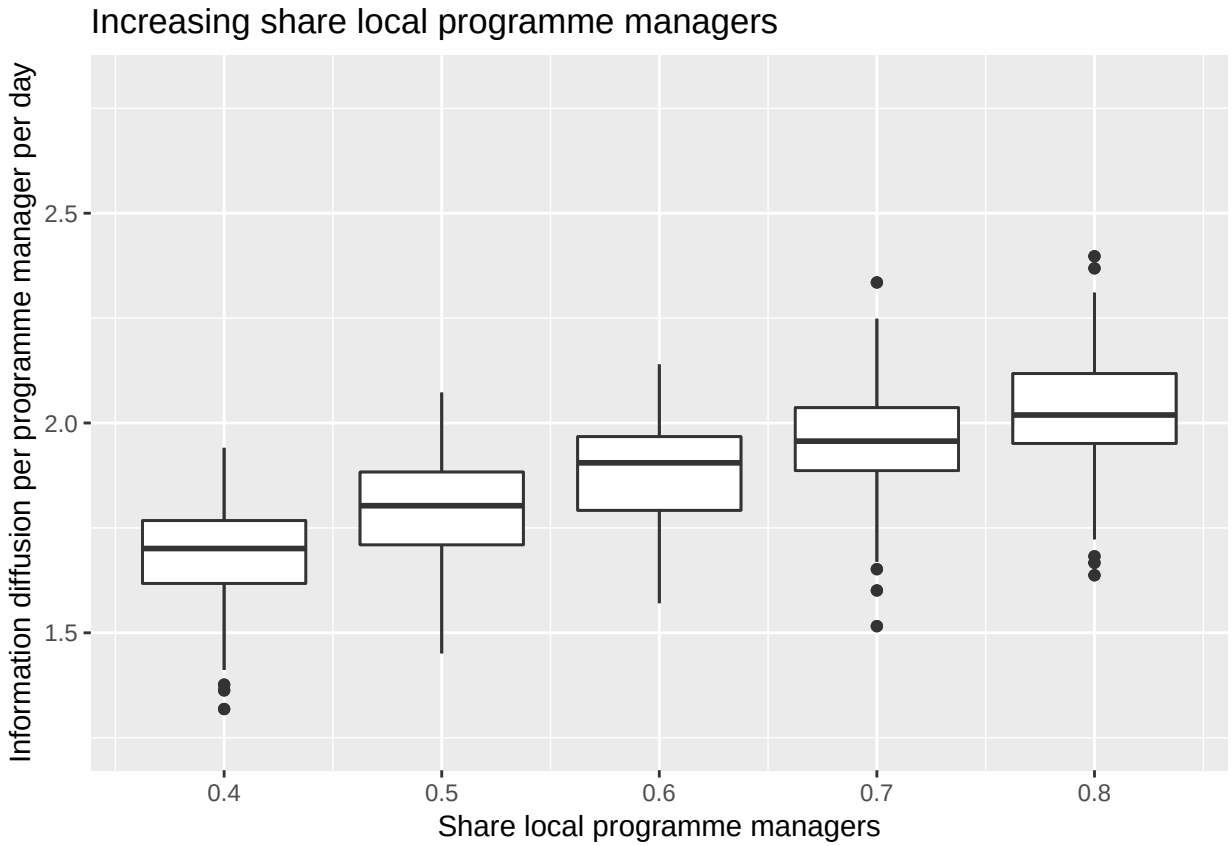

```
## Warning: Removed 6 rows containing non-finite values (stat_boxplot).
```

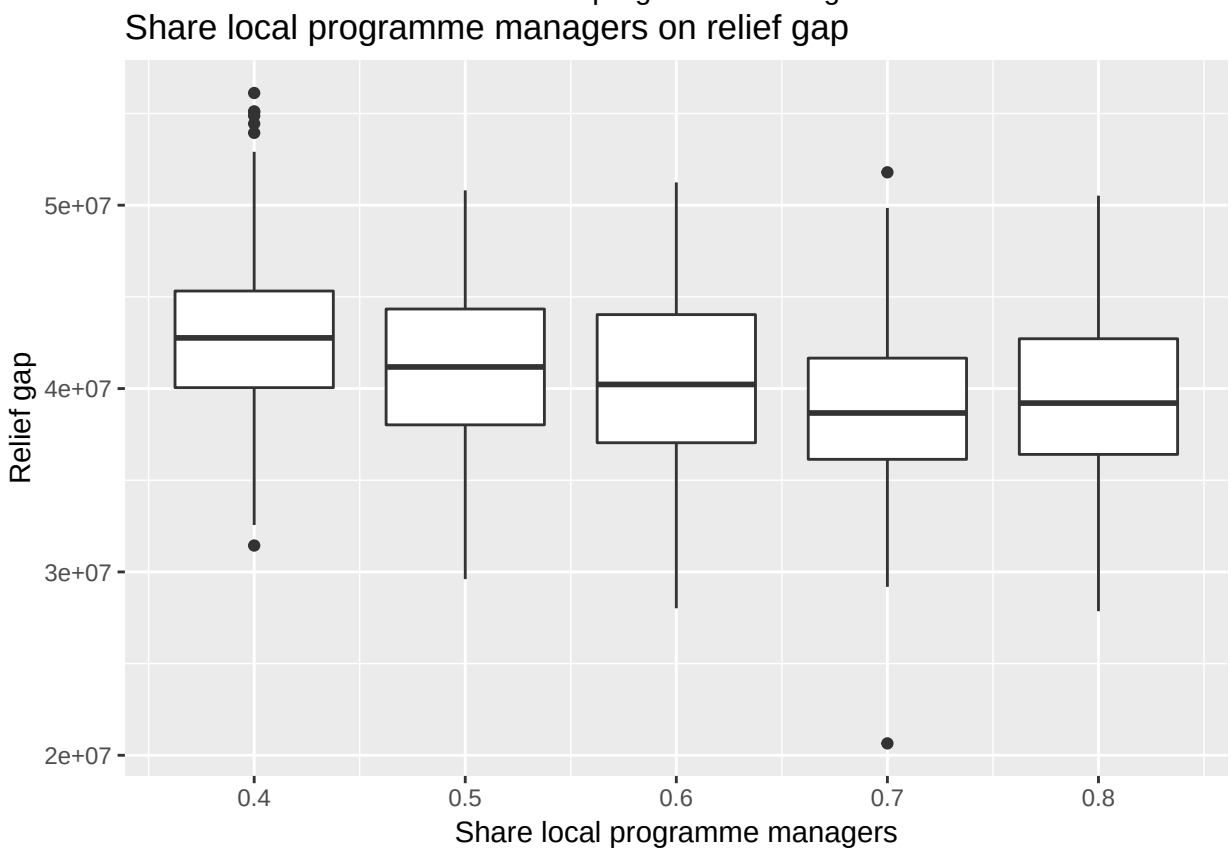


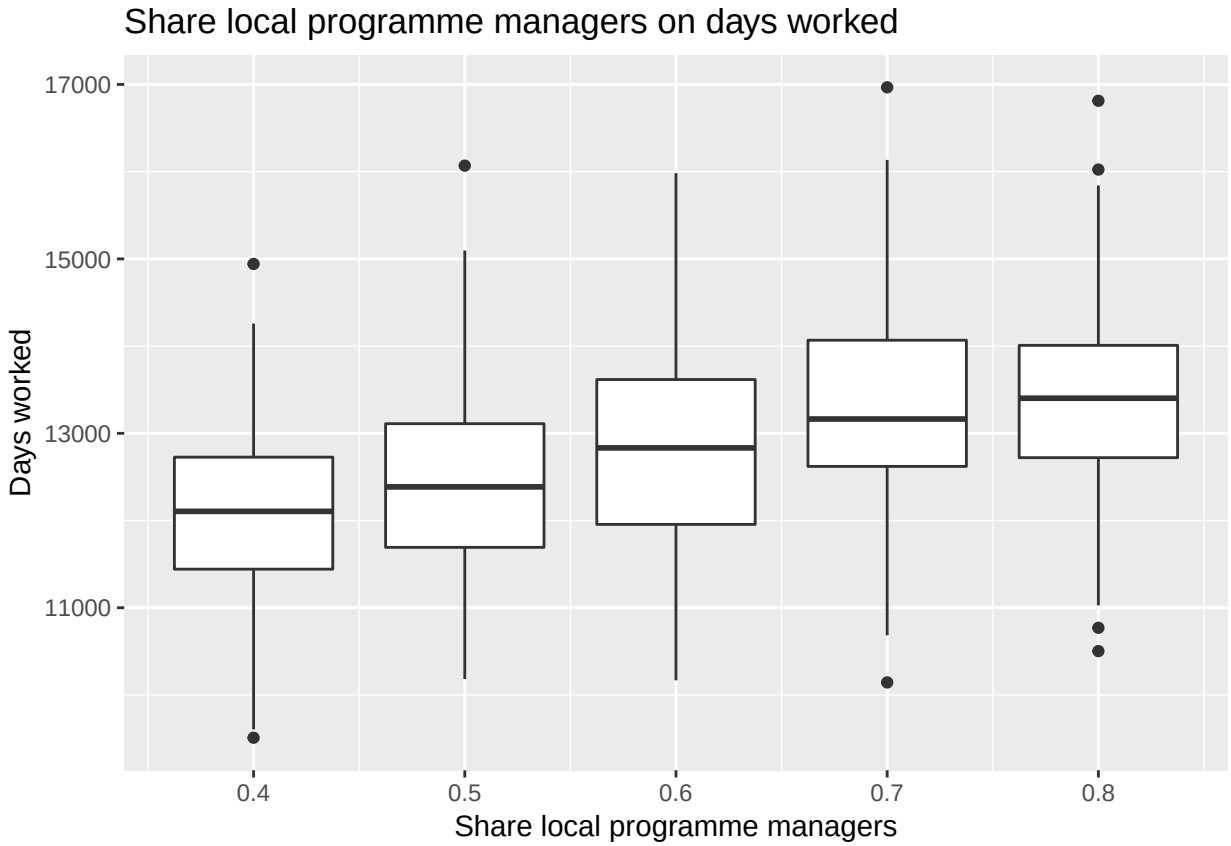

```
## Saving 6.5 x 4.5 in image
## Warning: Removed 93 rows containing non-finite values (stat_boxplot).
## Saving 6.5 x 4.5 in image
## Warning: Removed 6 rows containing non-finite values (stat_boxplot).
```

Share local - international

As the number of local programme mangers increases, the total information diffused increases but less as is the case as for both forms of increasing willigness to share information.But note 80% local programme mangers is more realistic that willigness to share 80% of all information. This increase could be due to the increase in days worked (that is shown in figure 2). The total gap decreases slightly as is shown in figure 3. The last 2 grapsh show the total information diffused and total gap corrected for to number of days worked. Not sure whether you are allowed to do this calculation (as the number was higher, and effects could also be indicrect?) Does this assume lineairty?


Increasing share local programme managers
Information diffusion per programme manager per day
2.5
2.0
1.5
0.4
0.5
0.6
0.7
0.8
Share local programme managers


Share local programme managers on relief gap
Relief gap
5e+07
4e+07
3e+07
2e+07
0.4
0.5
0.6
0.7
0.8
Share local programme managers

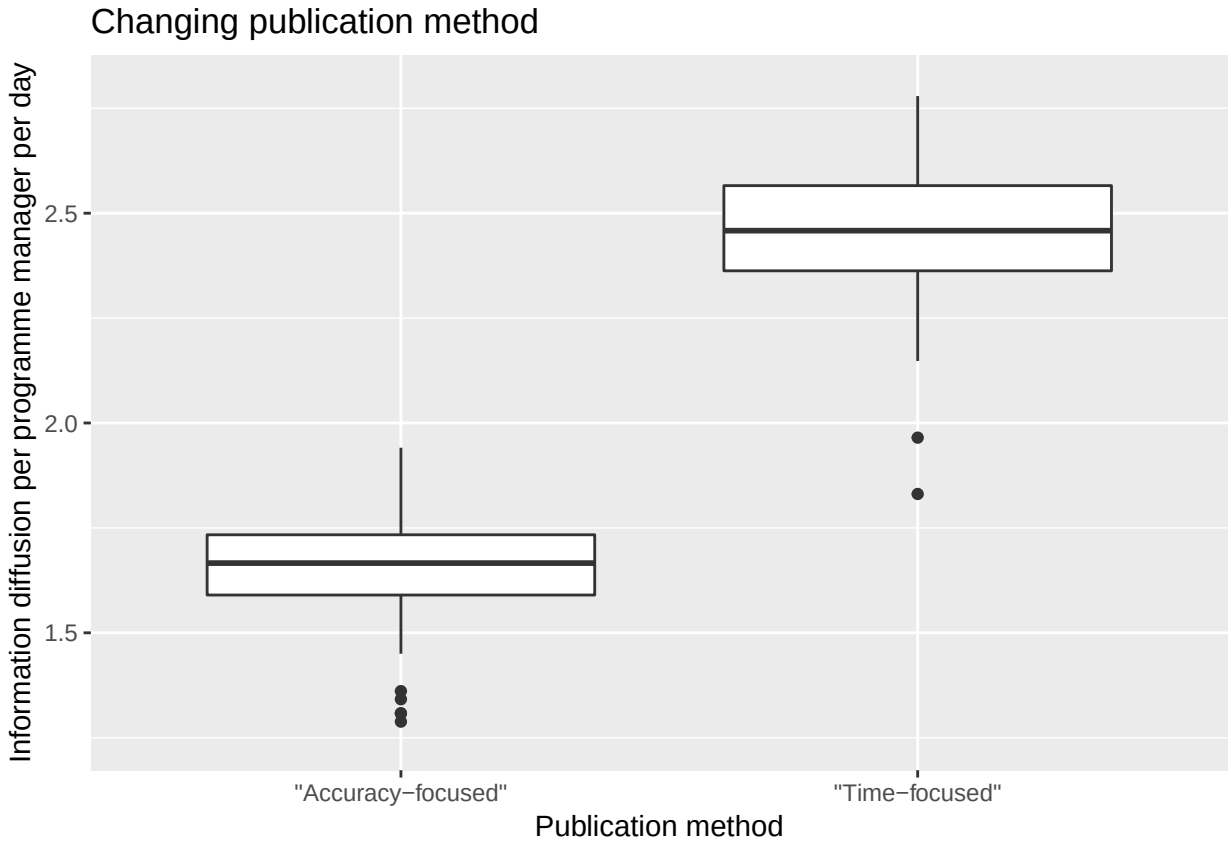


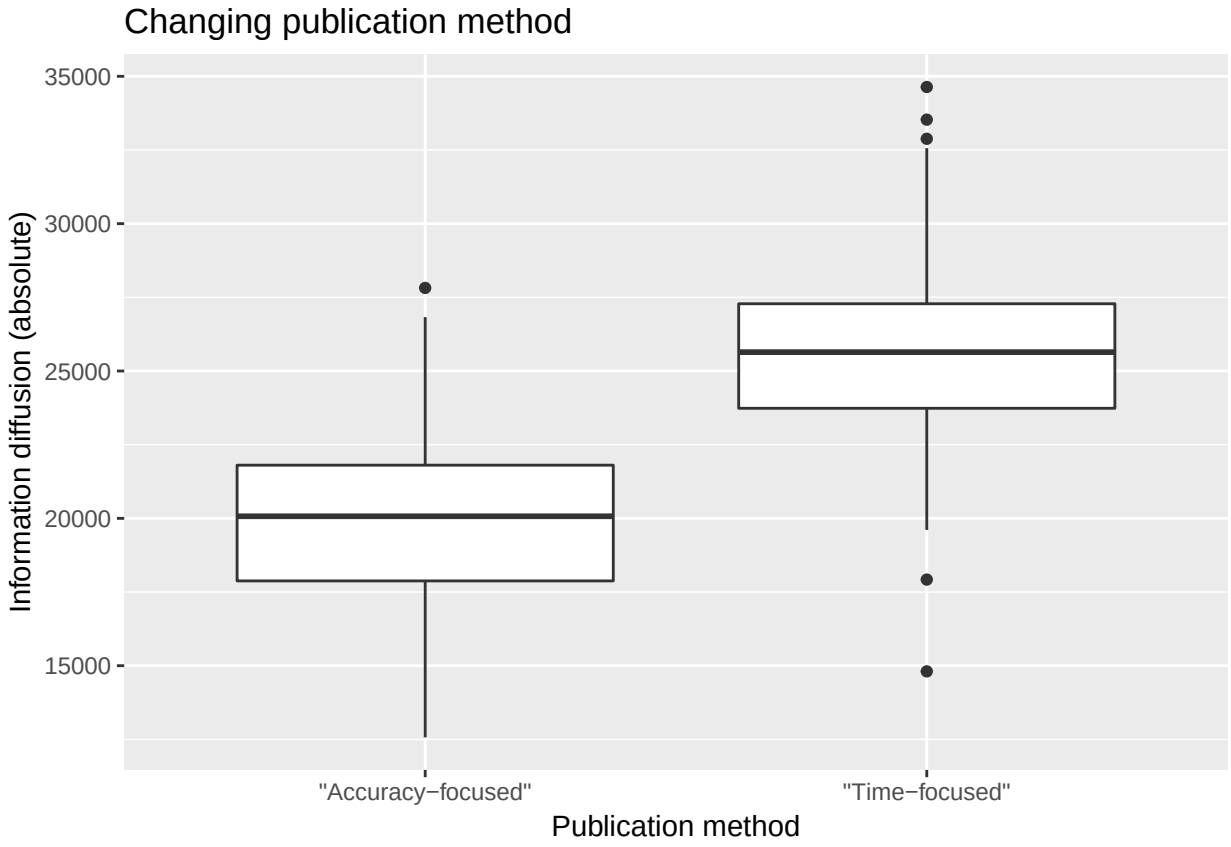

```
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
```

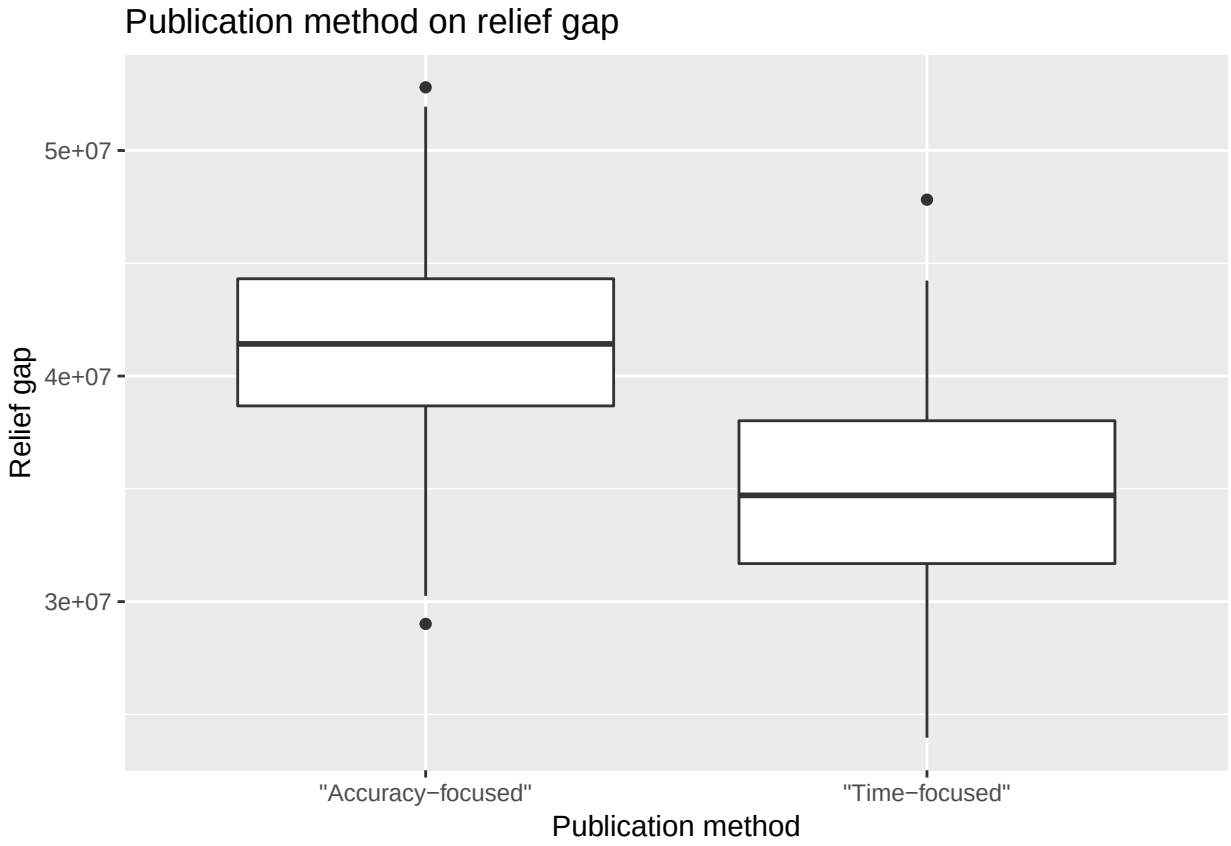
Publication method

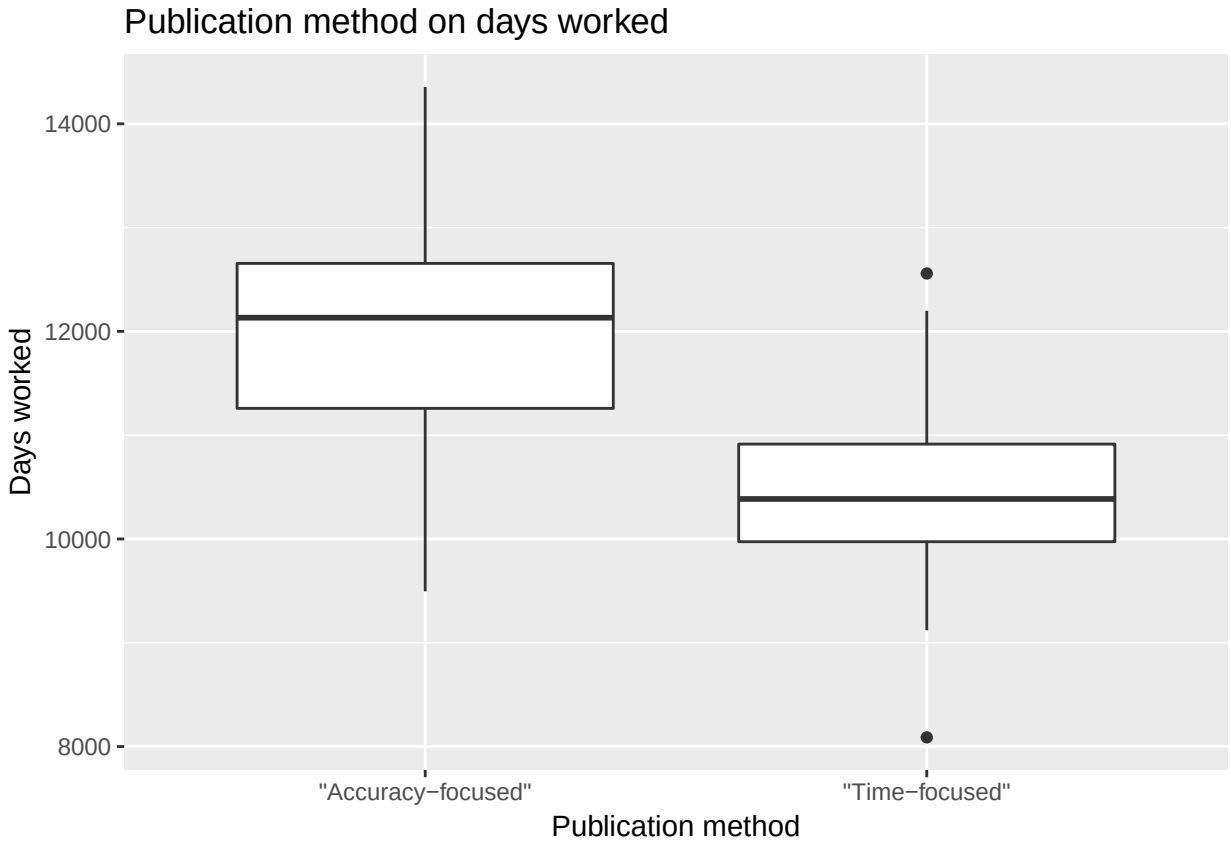
Publication strategies that are focused on fast publication of inaccuracte information sharing lead to more information diffusion (figure 1) and smaller relief gaps (figure 2) with less effort (lower number of days worked).

Did expect the effect to be so strong (it is the most effective strategie).


Changing publication method
Information diffusion per programme manager per day
2.5
2.0
1.5
"Accuracy−focused"
"Time−focused"
Publication method


Changing publication method
Information diffusion (absolute)
35000
30000
25000
20000
15000
"Accuracy−focused"
"Time−focused"
Publication method


Publication method on relief gap
Relief gap
5e+07
4e+07
3e+07
"Accuracy−focused"
"Time−focused"
Publication method

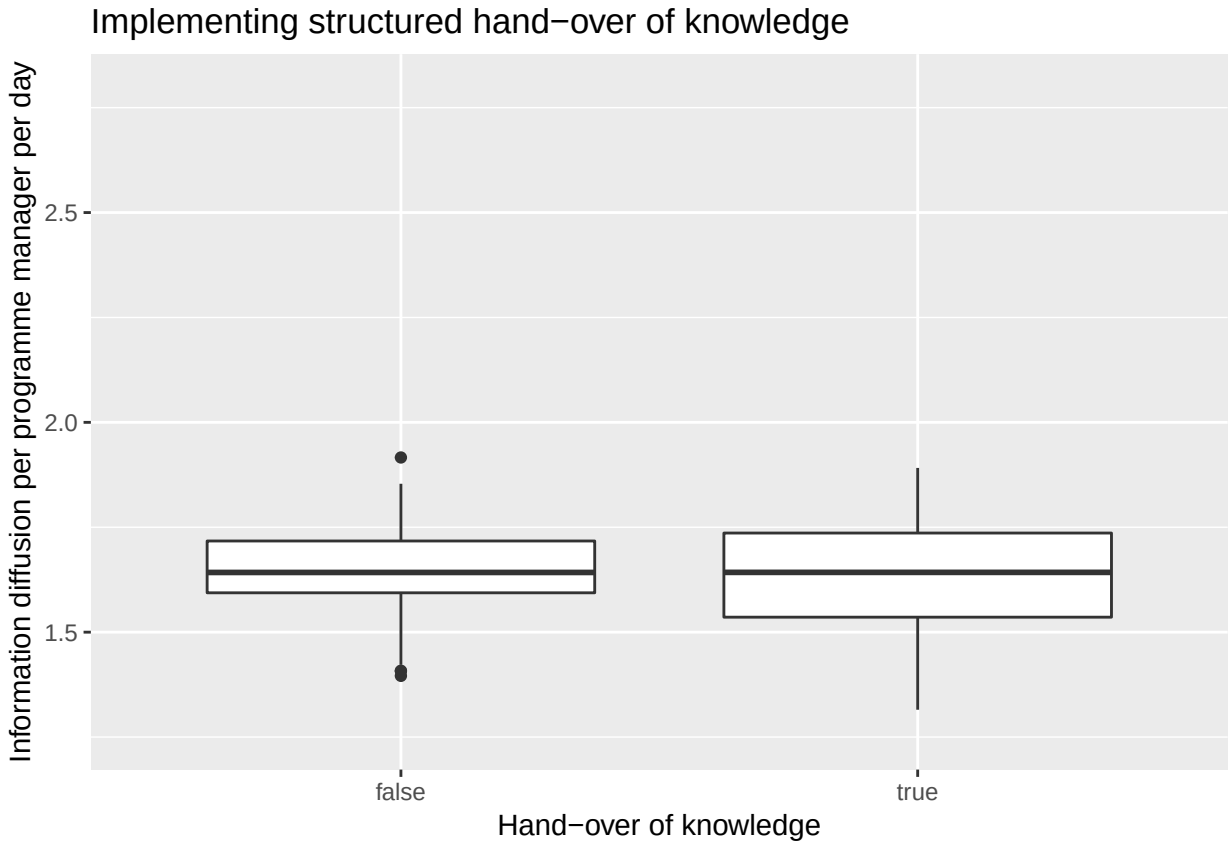

Publication method on days worked
Days worked
14000
12000
10000
8000
"Accuracy−focused"
"Time−focused"
Publication method

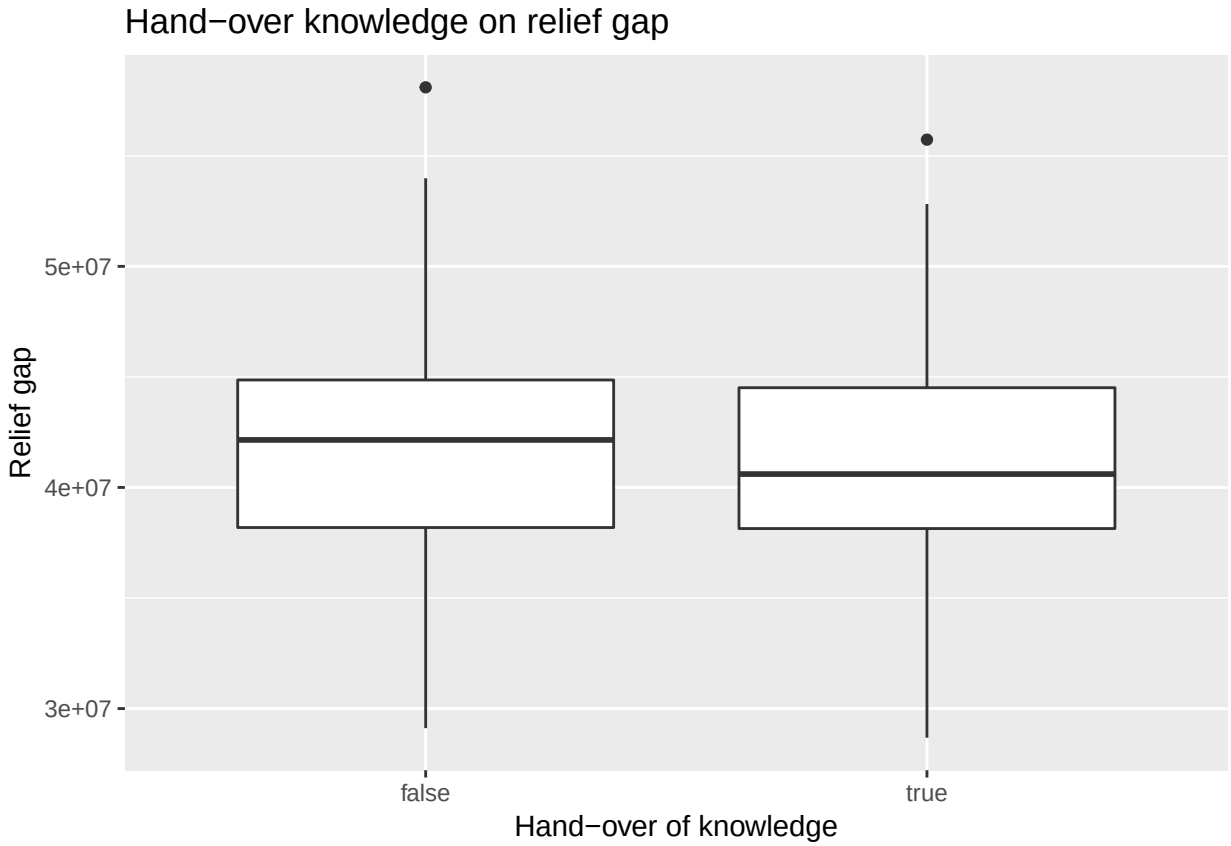
```
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
```

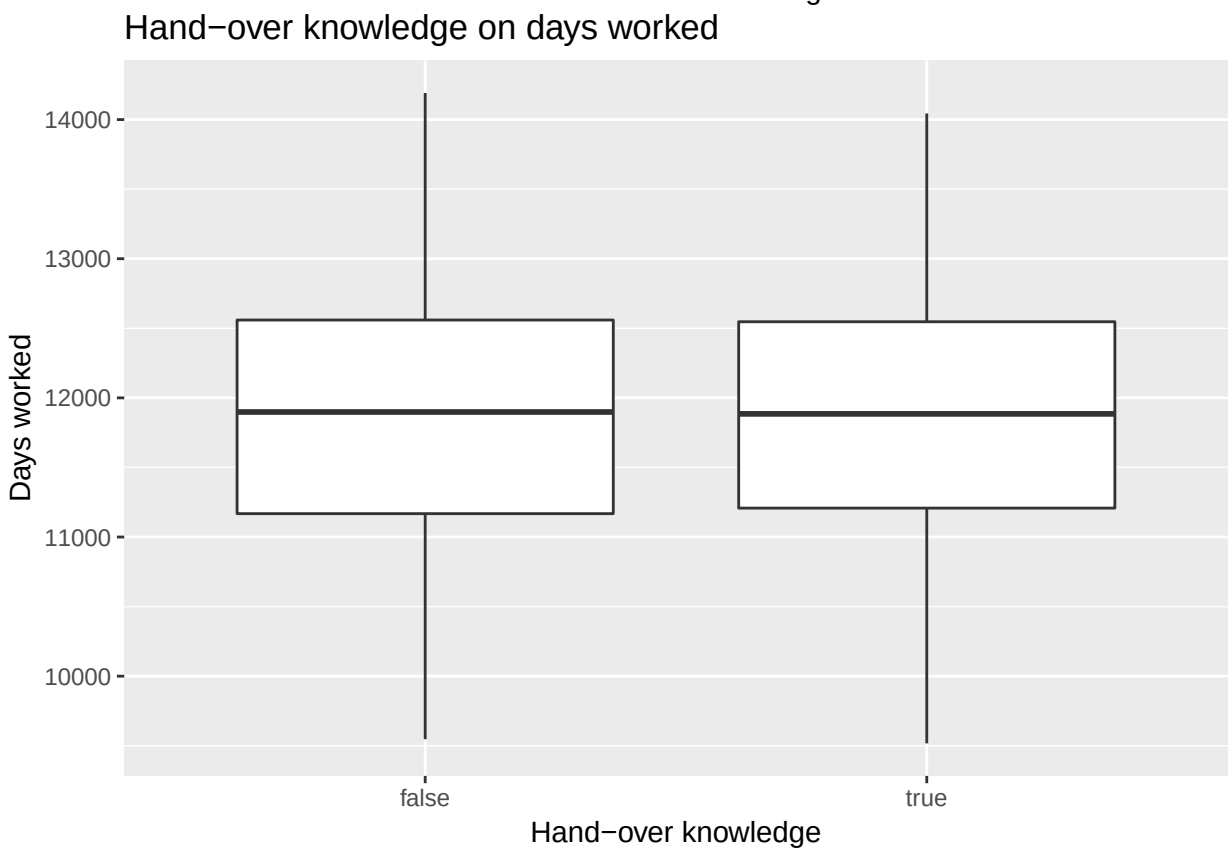
Hand over knowledge

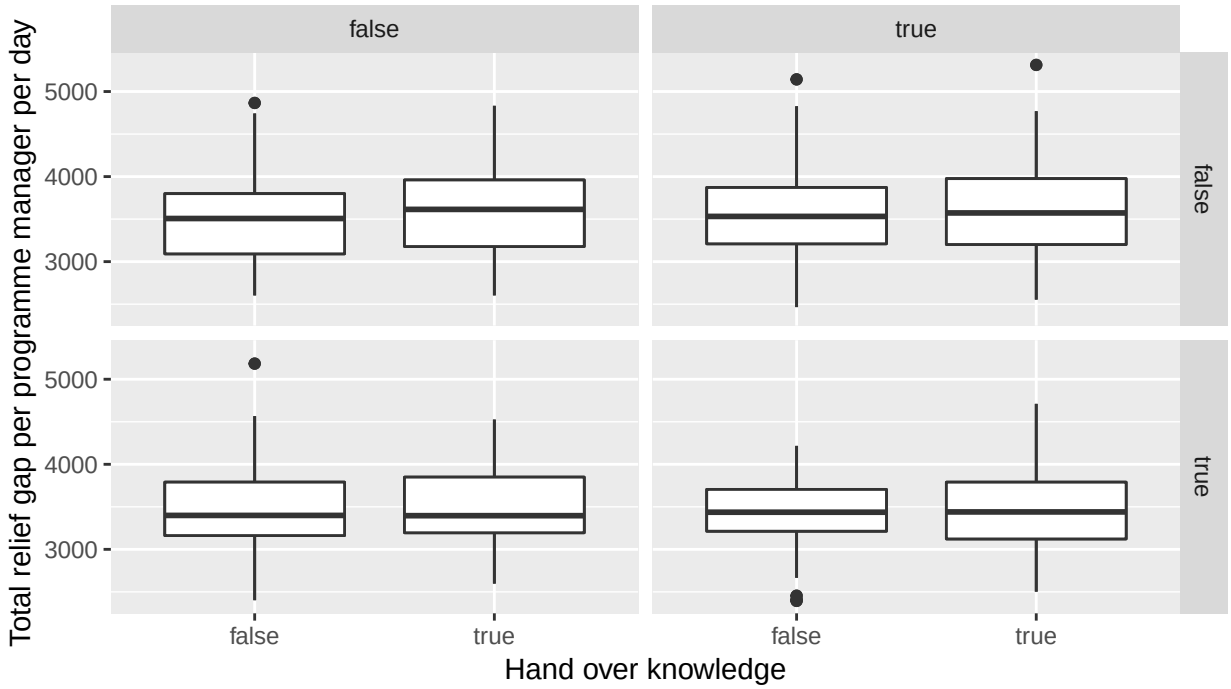
Focusing on the strategie hand over knowledge has no signification effect on information diffusion, the total relief gap or the total days worked.

```
## Warning: Removed 1 rows containing non-finite values (stat_boxplot).
```

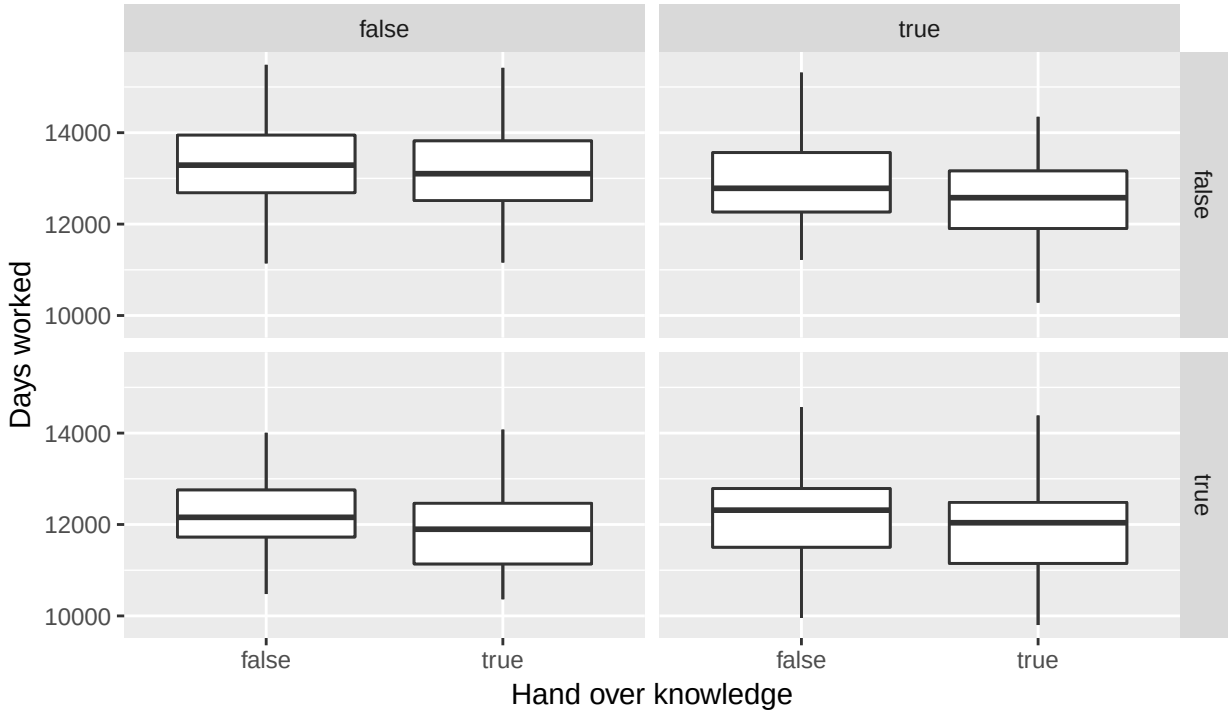



Hand−over knowledge on relief gap
Relief gap
5e+07
4e+07
3e+07
false
true
Hand−over of knowledge


Hand−over knowledge on days worked
Days worked
14000
13000
12000
11000
10000
false
true
Hand−over knowledge



Horizontal is social sharing. Vertical is IM sharing. In the case of no sharing (top left), there is an (weak) indication that handing over knowledge has an effect on the total relief gap.



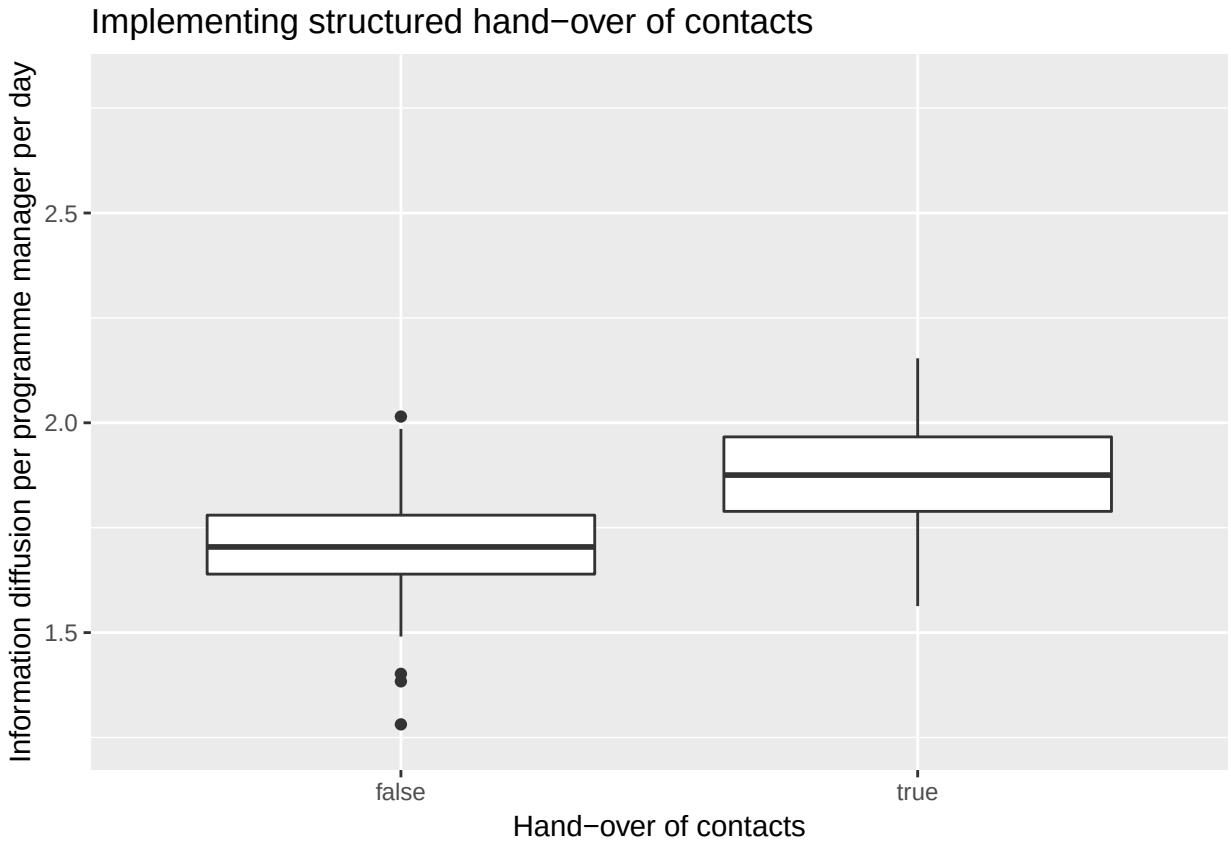
Horizontal is social sharing. Vertical is IM sharing. No effect on days worked.

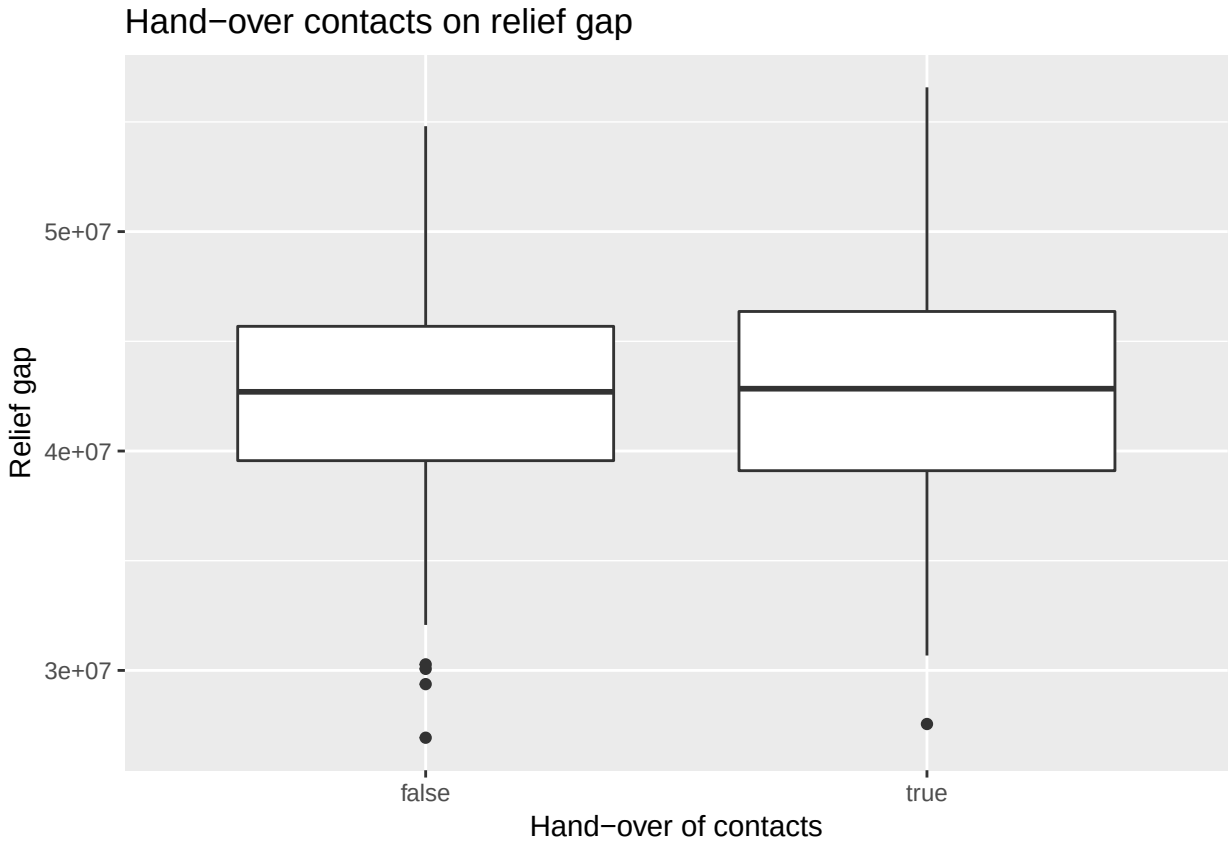
```
## Saving 6.5 x 4.5 in image
## Warning: Removed 1 rows containing non-finite values (stat_boxplot).
## Saving 6.5 x 4.5 in image
```

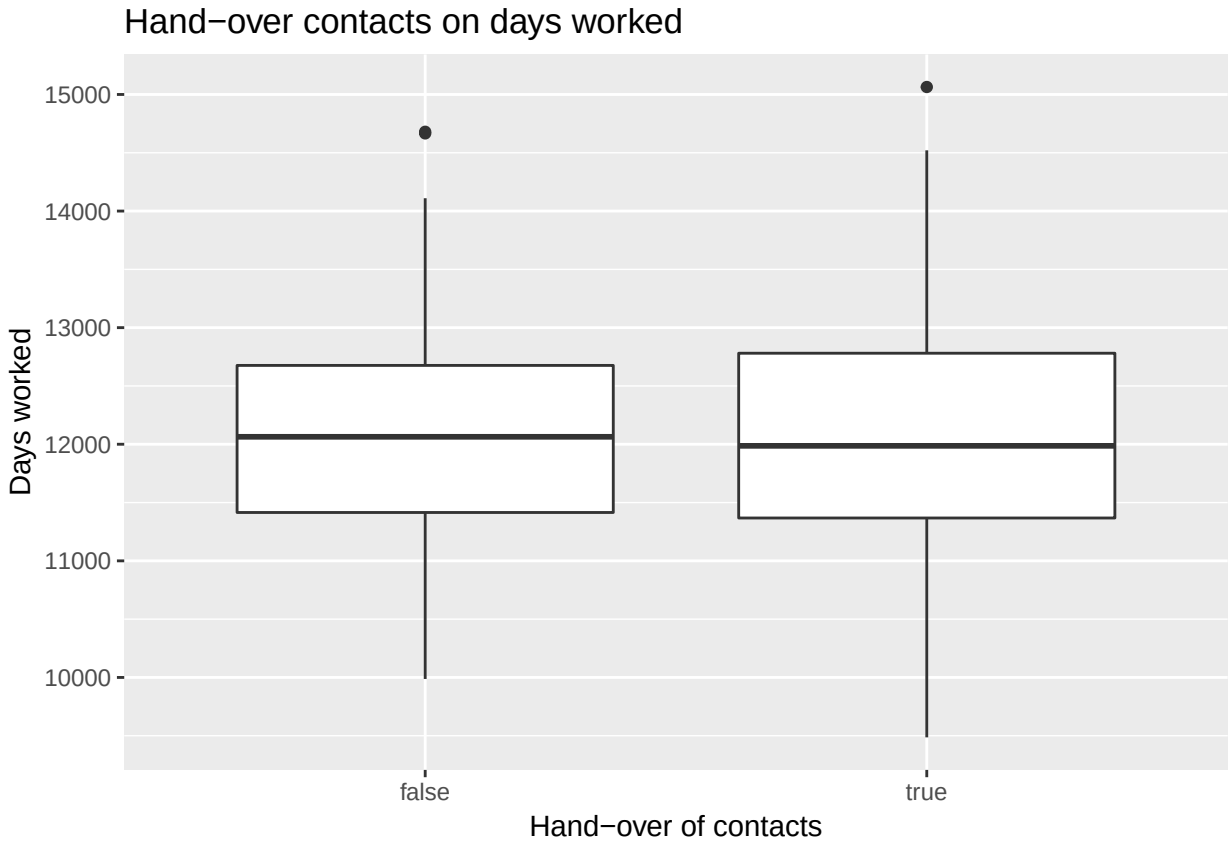
```
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
```

Hand over contacts

In contrast to focusing on handing over knowledge does the strategie hand over contacts slightly increase the information diffused and a small decrease of the relief gap and days worked.




Hand−over contacts on relief gap
Relief gap
5e+07
4e+07
3e+07
false
true
Hand−over of contacts

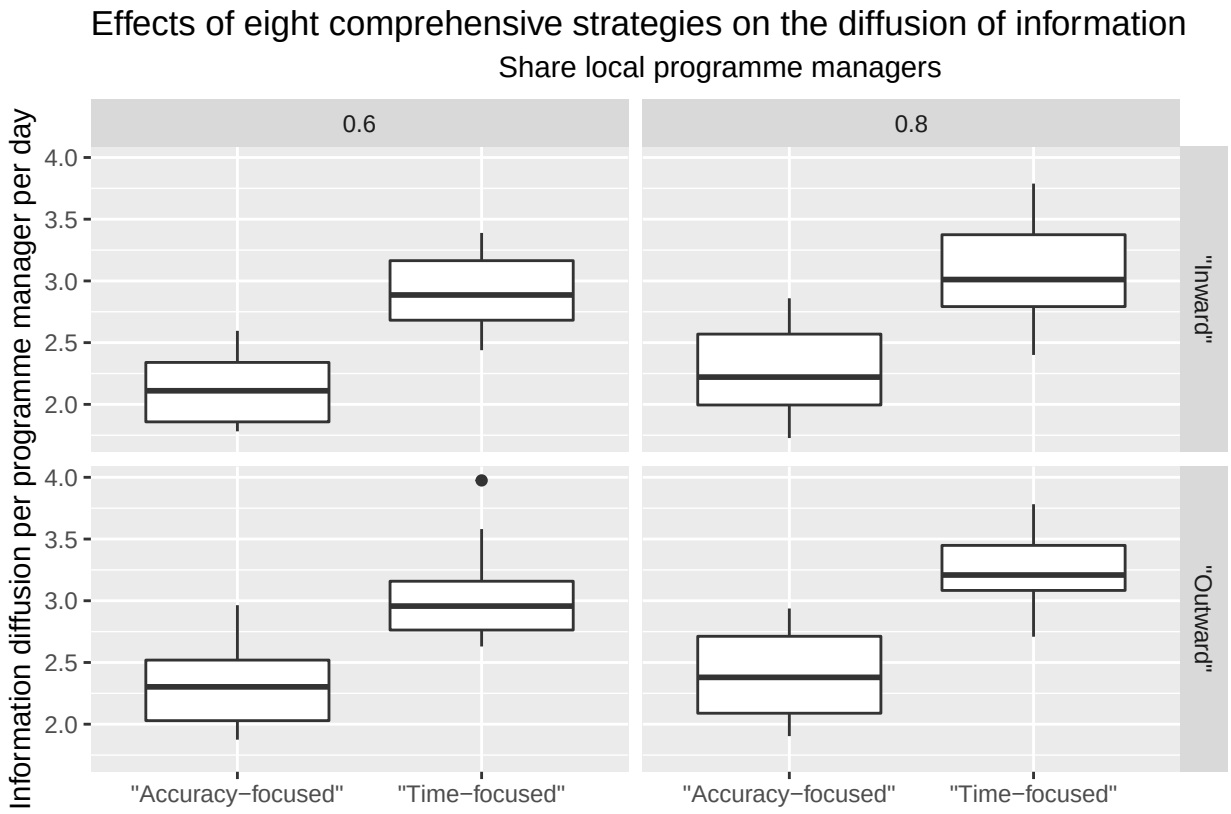

Hand−over contacts on days worked
Days worked
15000
14000
13000
12000
11000
10000
false
true
Hand−over of contacts

```
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
## Saving 6.5 x 4.5 in image
```

4. Analysis of comprehensive strategies

No supprising behavriour: Information diffusion is highest with a time-focused, local-based, high inter-organisations information sharing strategy.

```
comperhansive_str = sqldf("SELECT * FROM comperhansive_str WHERE shocks = 12")

boxplot_comperhansive_str = ggplot(comperhansive_str, aes(x=comperhansive_str$RS_publication_method, y=

print(boxplot_comperhansive_str)
```

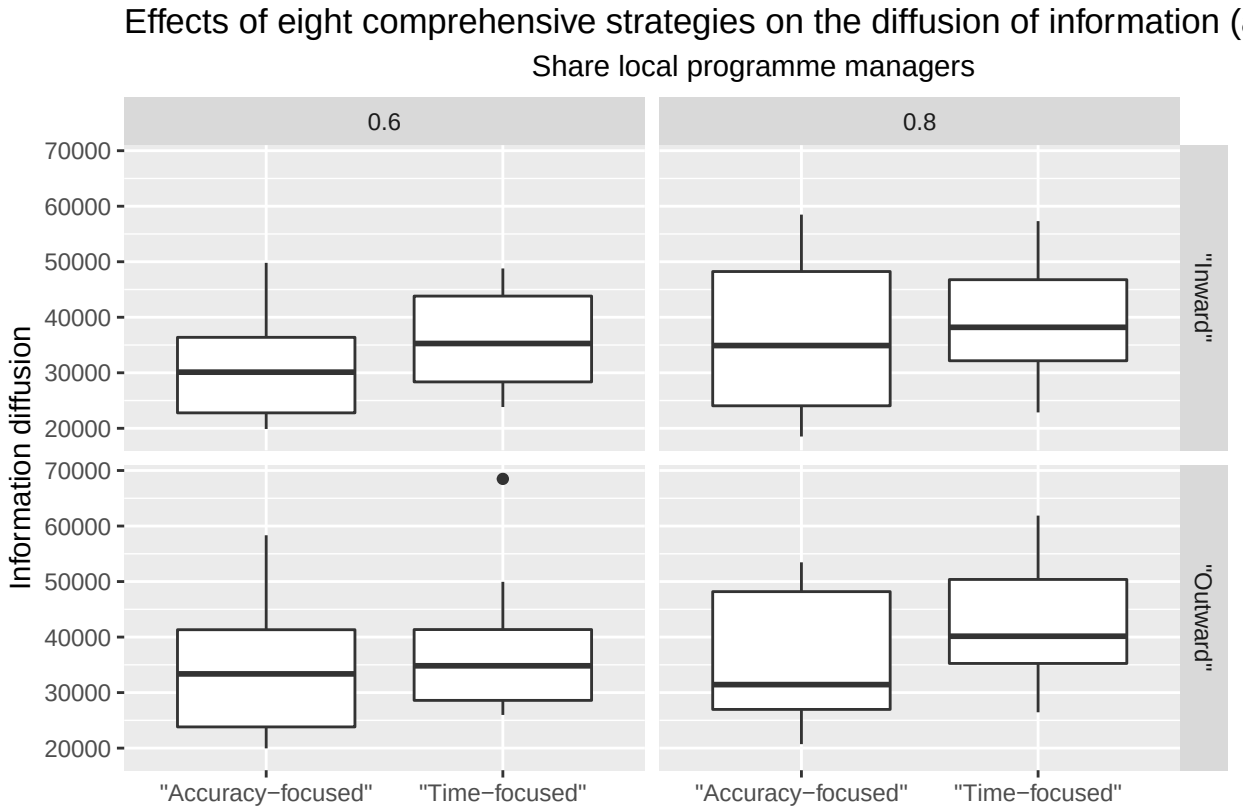



```
ggsave("plots/3_1boxplot_comperhansive_str.png",boxplot_comperhansive_str)
```

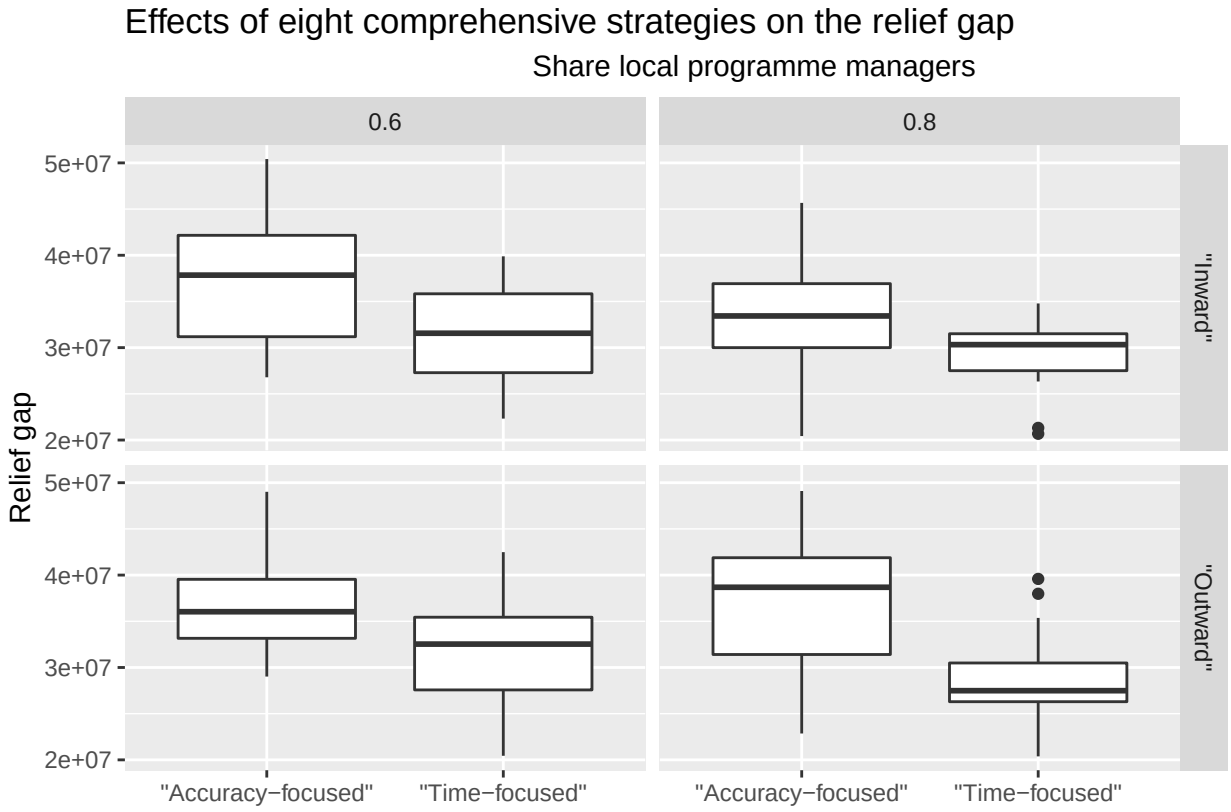
```
## Saving 6.5 x 4.5 in image
```

Comperhansive strategies on the relief gap and days worked KPI

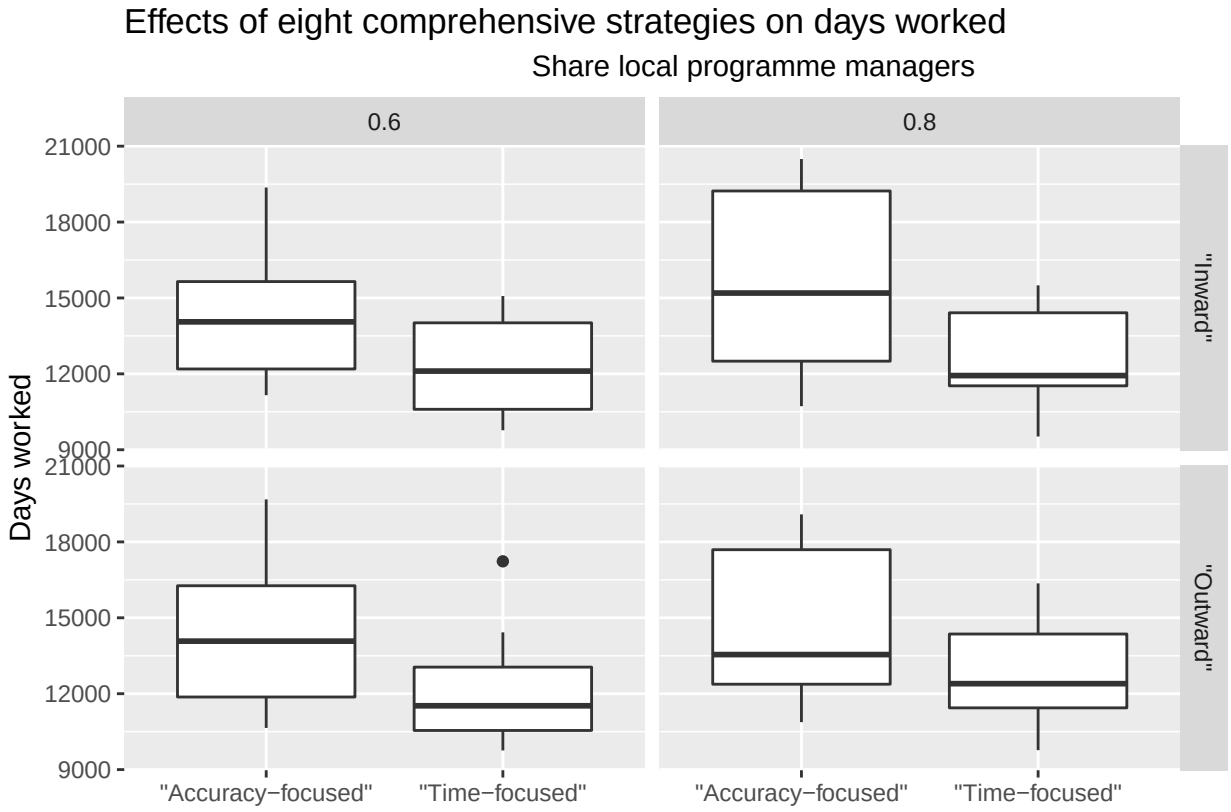
```
# Not in thesis
boxplot_comperhansive_str_inf_diff_absolute = ggplot(comperhansive_str, aes(x=comperhansive_str$RS_publi
       x = "", y = "Information diffusion" )

print(boxplot_comperhansive_str_inf_diff_absolute)
```




```
boxplot_comperhansive_str_gap = ggplot(comperhansive_str, aes(x=comperhansive_str$RS_publication_method
       x = "", y = "Relief gap" )

print(boxplot_comperhansive_str_gap)
```




```
boxplot_comperhansive_str_days = ggplot(comperhansive_str, aes(x=comperhansive_str$RS_publication_metho
       x = "", y = "Days worked" )

print(boxplot_comperhansive_str_days)
```




```
ggsave("plots/3_2boxplot_comperhansive_str_gap.png",boxplot_comperhansive_str_gap)
```

```
## Saving 6.5 x 4.5 in image
```

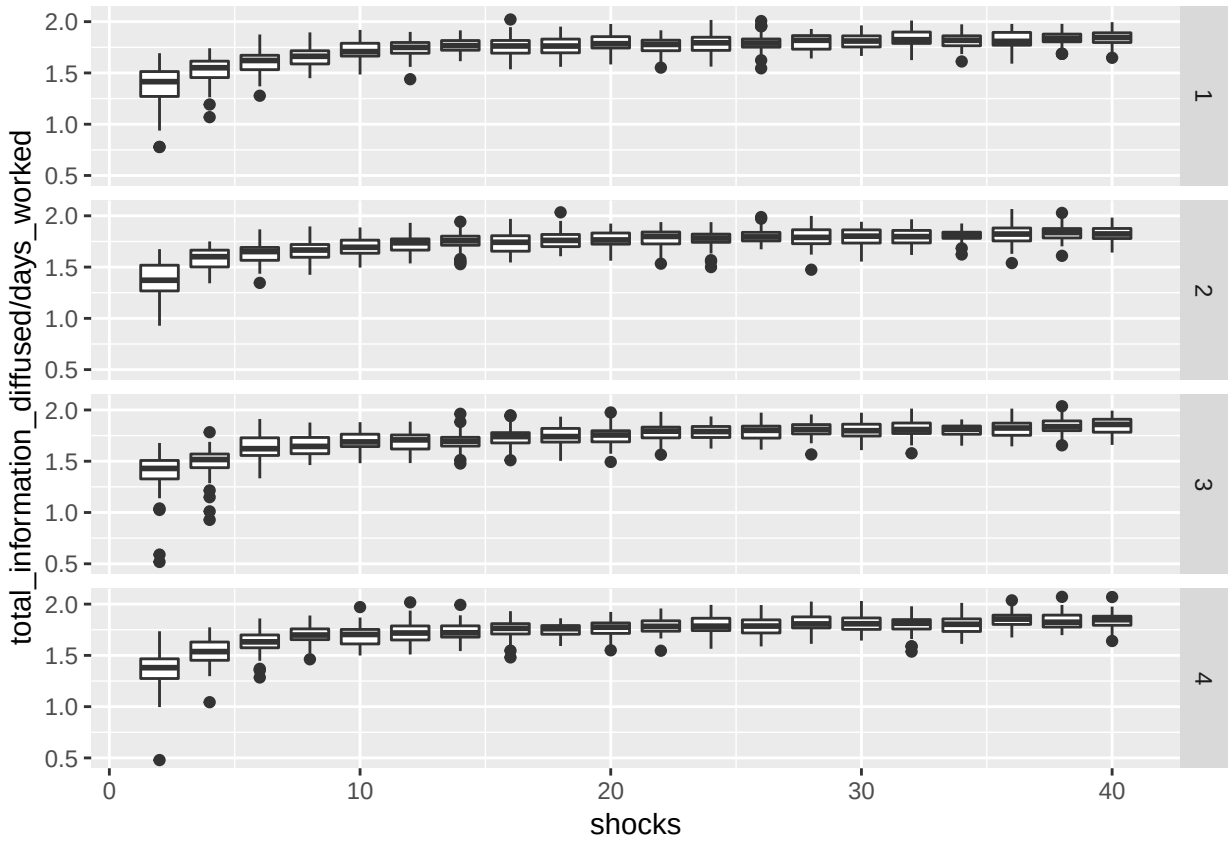
```
ggsave("plots/3_3boxplot_comperhansive_str_days.png",boxplot_comperhansive_str_days)
```

```
## Saving 6.5 x 4.5 in image
```

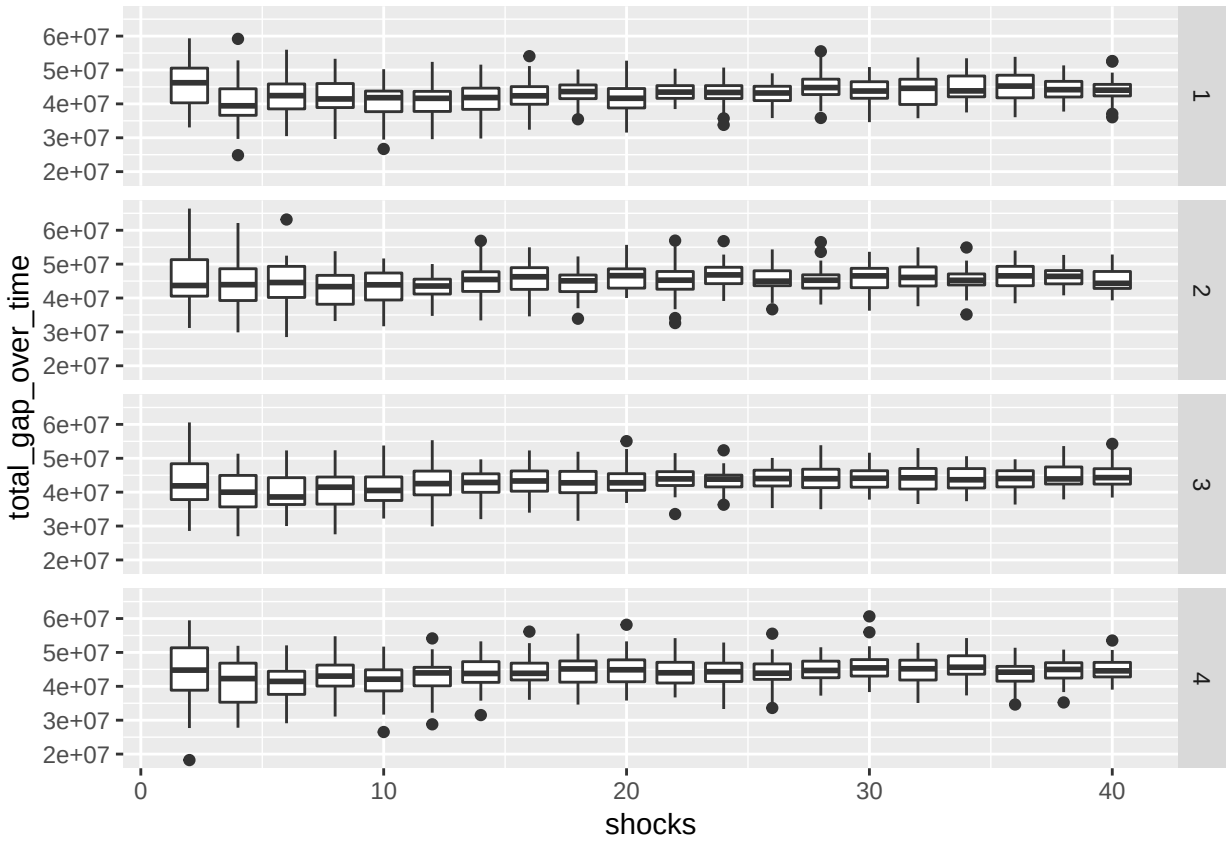
5. Analayis for structural validation

Changing the accuracy assumption - Effect on KPIs for various numbers of shocks

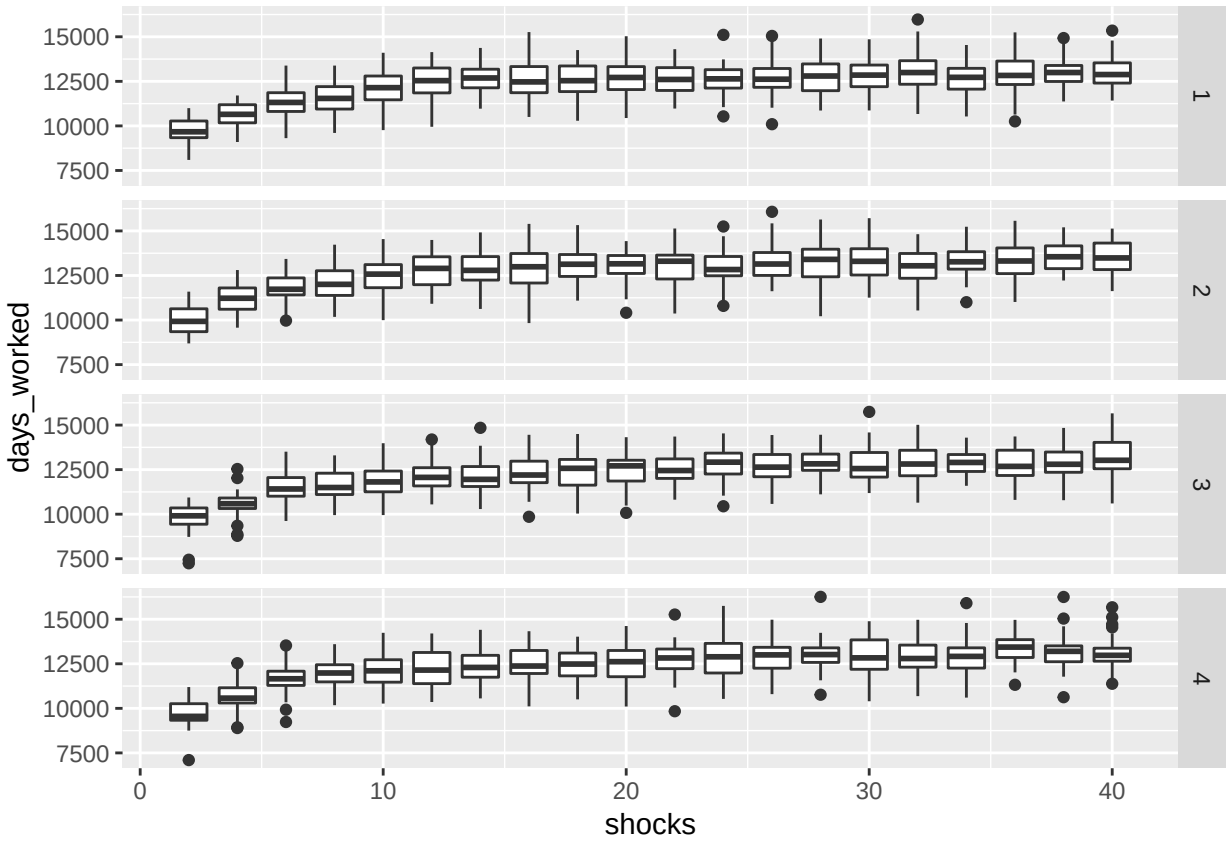
```
boxplot_effect_distributions_shocks_inf_diff = ggplot(data=effect_distributions_shocks, aes(x=shocks, y=

print(boxplot_effect_distributions_shocks_inf_diff)
```




```
ggsave("plots/4_2boxplot_effect_distributions_shocks_inf_diff.png",boxplot_effect_distributions_shocks_

boxplot_effect_distributions_shocks_gap = ggplot(data=effect_distributions_shocks, aes(x=shocks, y=total

print(boxplot_effect_distributions_shocks_gap)
```



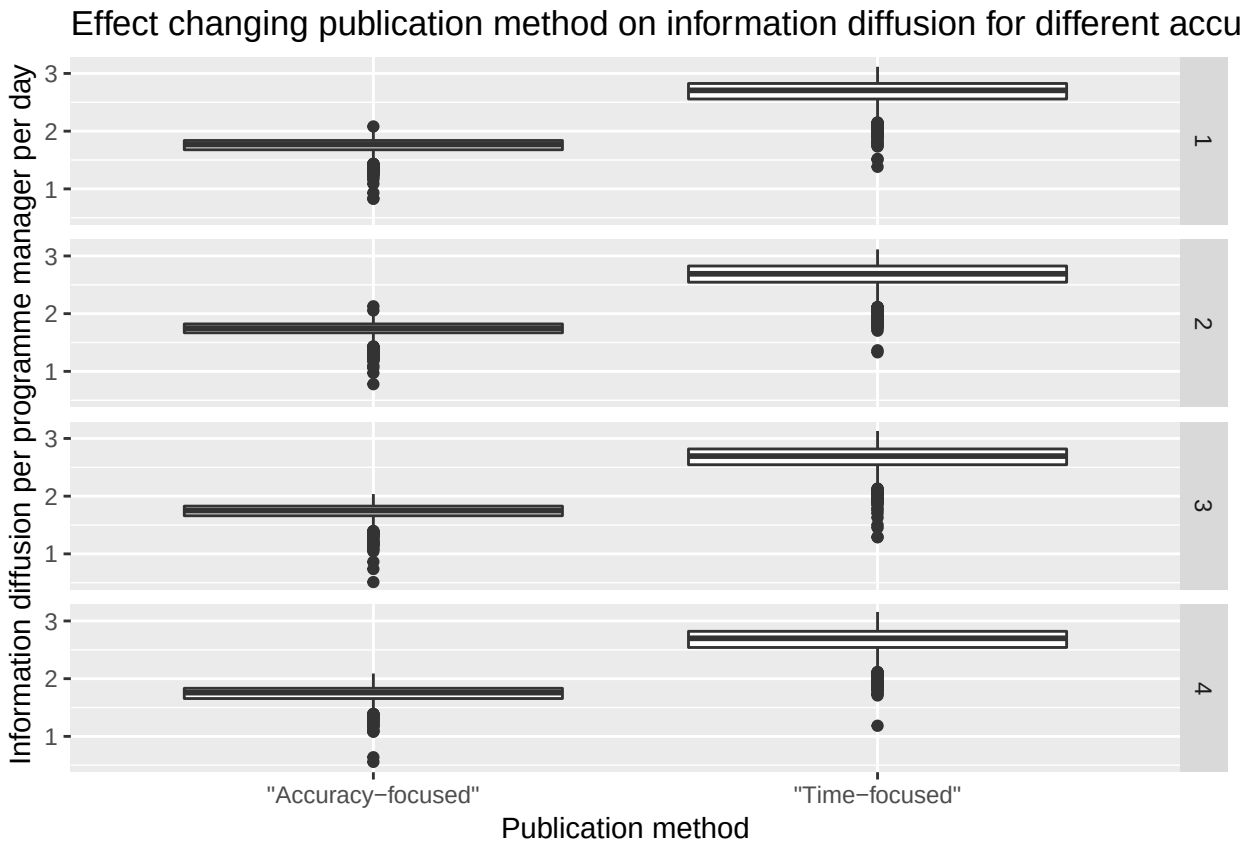

```
ggsave("plots/4_3boxplot_effect_distributions_shocks_gap.png",boxplot_effect_distributions_shocks_gap,w

boxplot_effect_distributions_shocks_days = ggplot(data=effect_distributions_shocks, aes(x=shocks, y= da

print(boxplot_effect_distributions_shocks_days)
```



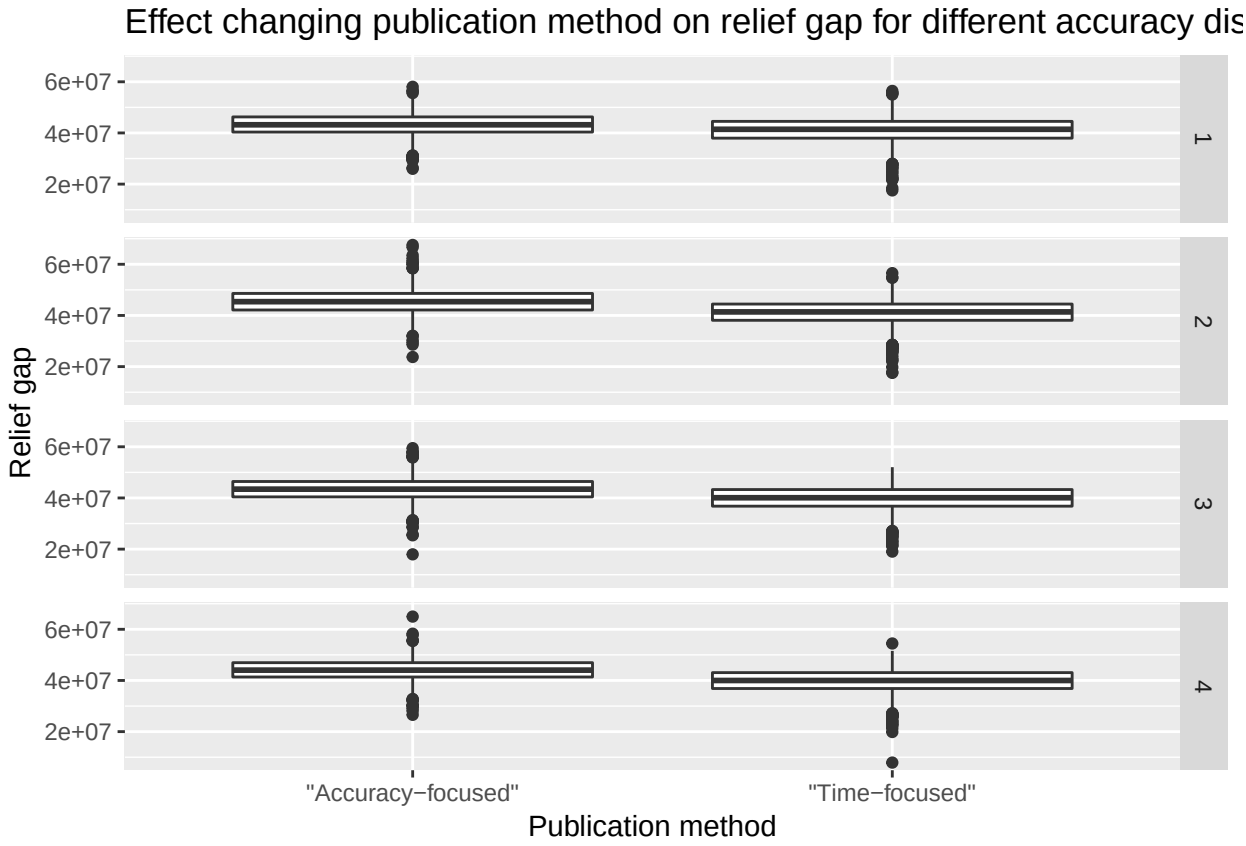

```
ggsave("plots/4_4boxplot_effect_distributions_shocks_days .png",boxplot_effect_distributions_shocks_days
```

Changing the accuracy assumption - Effect on KPIs for accuracy focused and time focused publication methods

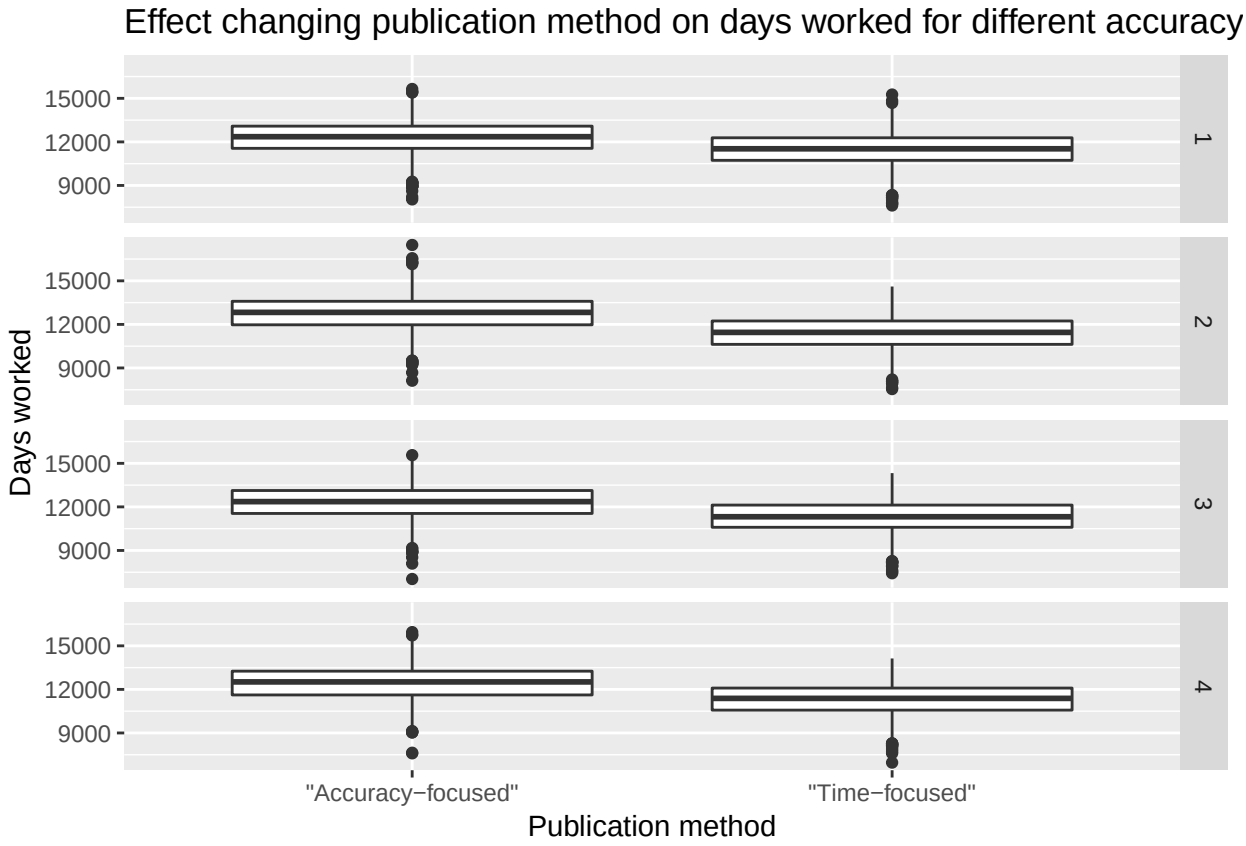
```
boxplot_Ind_str_publication_method_distribution_inf_diff = ggplot(Ind_str_publication_method_distributi
  geom_boxplot() +
 labs(title = "Effect changing publication method on information diffusion for different accuracy distri
       x = "Publication method", y = "Information diffusion per programme manager per day")

print(boxplot_Ind_str_publication_method_distribution_inf_diff)
```




```
ggsave("plots/4_5boxplot_Ind_str_publication_method_distribution_inf_diff.png", boxplot_Ind_str_publica


boxplot_Ind_str_publication_method_distribution_gap = ggplot(Ind_str_publication_method_distribution, a
  geom_boxplot() +
 labs(title = "Effect changing publication method on relief gap for different accuracy distributions",
       x = "Publication method", y = "Relief gap")

print(boxplot_Ind_str_publication_method_distribution_gap)
```




```
ggsave("plots/4_6boxplot_Ind_str_publication_method_distribution_gap.png", boxplot_Ind_str_publication_


boxplot_Ind_str_publication_method_distribution_days = ggplot(Ind_str_publication_method_distribution, a
  geom_boxplot() +
 labs(title = "Effect changing publication method on days worked for different accuracy distributions",
       x = "Publication method", y = "Days worked")

print(boxplot_Ind_str_publication_method_distribution_days)
```

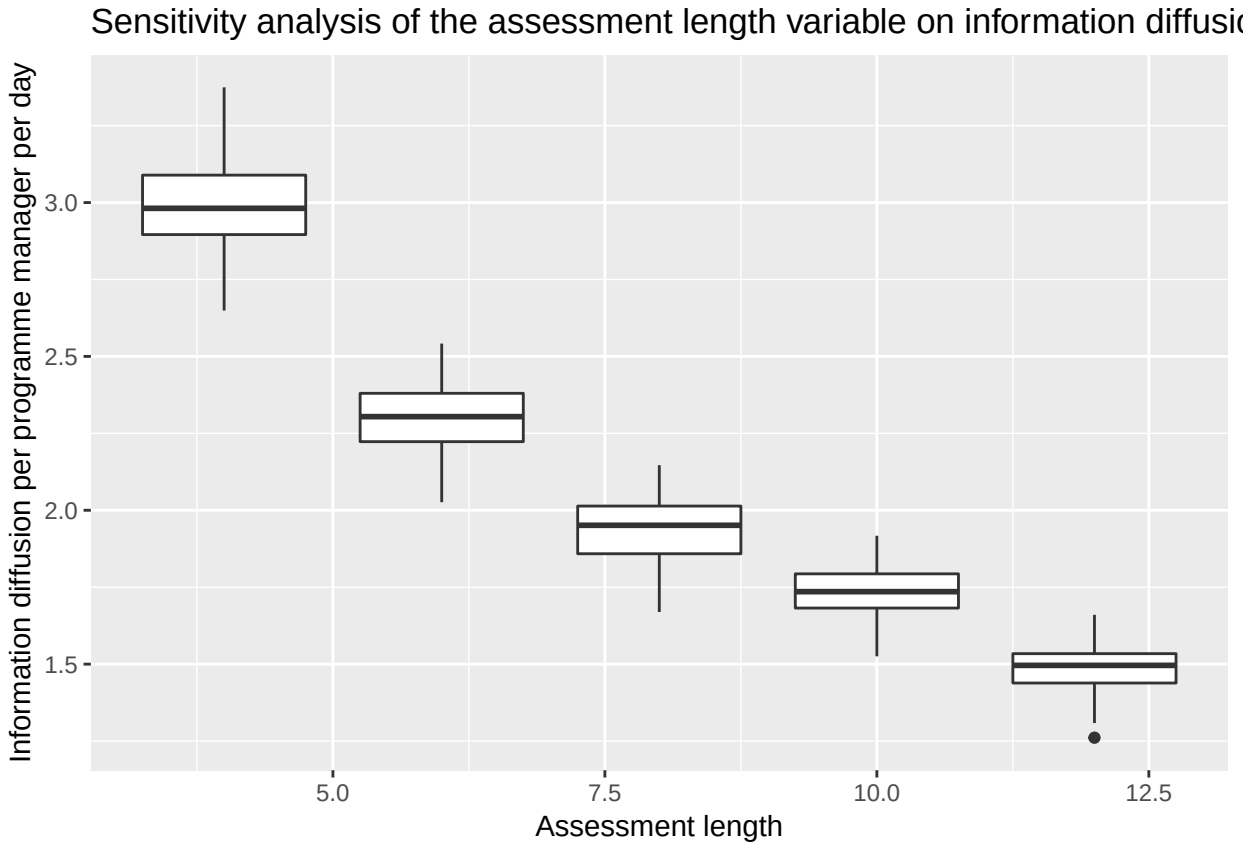



```
ggsave("plots/4_7boxplot_Ind_str_publication_method_distribution_days.png", boxplot_Ind_str_publication_
```

Sensitivity analysis on assessment length

It is problematic that the number of assessments increases (but the accuracy stays the same -> benefit of increasing speed).

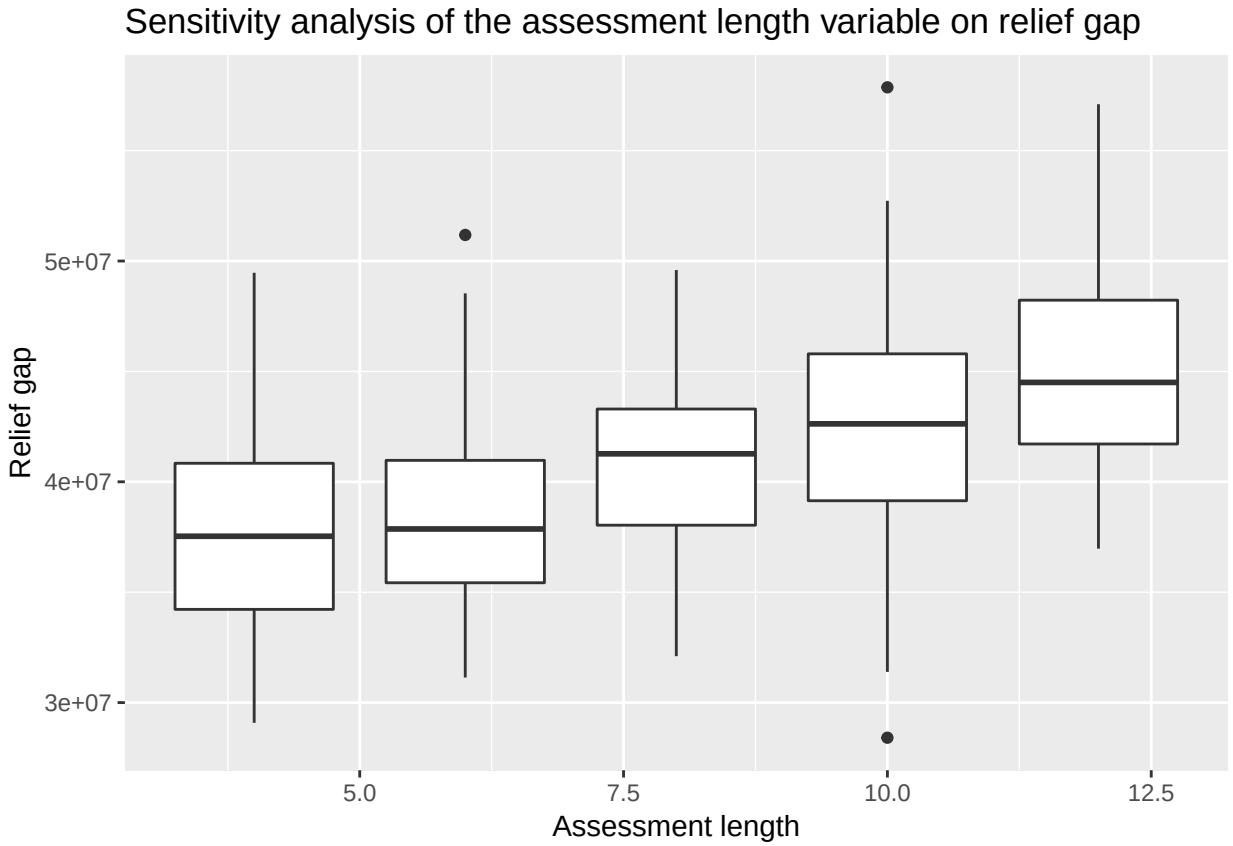
```
boxplot_sensitivity_assessment_length_inf_diff = ggplot(sensitivity_assessment_length, aes(x=assessment_

print(boxplot_sensitivity_assessment_length_inf_diff)
```




```
ggsave("plots/5_1boxplot_sensitivity_inf_diff.png",boxplot_sensitivity_assessment_length_inf_diff)
```

```
## Saving 6.5 x 4.5 in image
```

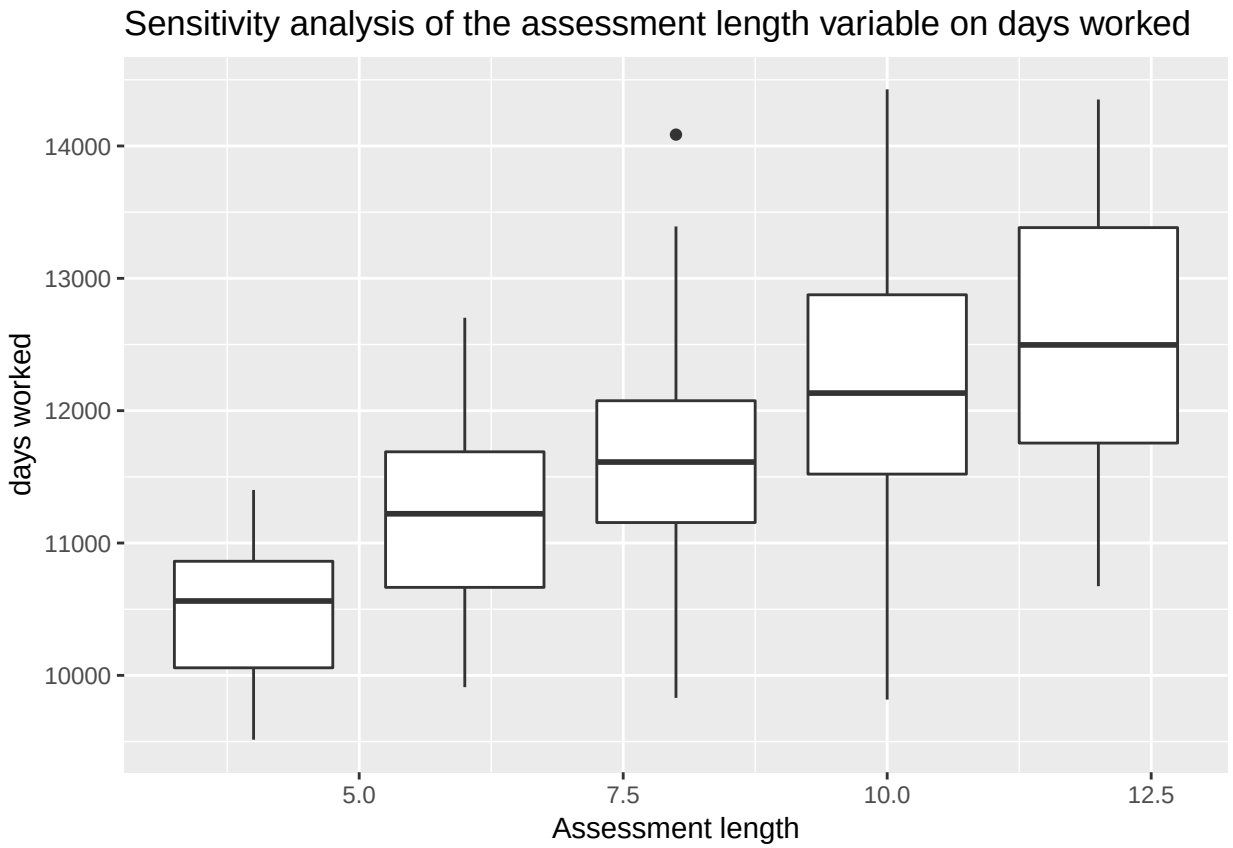
```
boxplot_sensitivity_assessment_length_gap = ggplot(sensitivity_assessment_length, aes(x=assessment_leng
       x = "Assessment length", y = "Relief gap")

print(boxplot_sensitivity_assessment_length_gap)
```




```
ggsave("plots/5_2boxplot_sensitivity_gap.png",boxplot_sensitivity_assessment_length_gap)
```

```
## Saving 6.5 x 4.5 in image
```

```
boxplot_sensitivity_assessment_length_days = ggplot(sensitivity_assessment_length, aes(x=assessment_leng
       x = "Assessment length", y = "days worked")

print(boxplot_sensitivity_assessment_length_days)
```




```
ggsave("plots/5_3boxplot_sensitivity_days.png",boxplot_sensitivity_assessment_length_days)
```

```
## Saving 6.5 x 4.5 in image
```