**Apêndice A –** **EXPERIMENTOS**

## planejamento

## AMEAÇAS A VALIDADE

### AMEAÇA A VALIDADE INTERNA

* A principio nossos experimentos serão realizados controlados por nós, provocando uma ameaça. Como contorno podemos usar exemplos do livro do Fowler para a construção do gabarito.
* (A interface gráfica amigável pode influenciar a pessoa que esta avaliando, caso outra pessoa avalie?)
* Sim, e essa é a ideia, né? Que influencie positivamente. De certa  
  forma, ao usar uma métrica objetiva, estaremos capturando essa  
  influência. Mas de fato, quando outra pessoa usa ela pode argumentar  
  coisas como "satisfação" que não são fáceis de medir objetivamente.
* (A utilização massiva de ferramentas existentes poderia causar uma ameaça, devido a aceitação, adaptação e entendimento do usuário em relação a nossa ferramenta -- Caso outra pessoa avalie)
* Comparar  o IDIFF com ferramentas capazes de detectar movimentações para não gerar a comparação que possa prover beneficios para o IDIFF

*Como utilizar métricas na comparação de diretórios? diff entre classes? como avaliar?*   
  
 Pois é... poderia pensar em precisão e cobertura para o caso dos  
diretórios, mas como ninguém mais faz isso, será estranha a  
comparação. Nesse ponto um experimento controlado com humanos teria  
suas vantagens, pois passaríamos um cenário (por exemplo, duas versões  
de código) e perguntaríamos o que houve entre essas duas versões em  
termos de refatoração. Aí poderíamos medir a corretude da resposta e o  
tempo que a pessoa levou para responder.

* > Provavelmente não faremos experimentos em grandes projetos, o que deixa a  
  > dúvida se o resultado também será satisfatório. Gerando limitação na  
  > generalização do resultado
* Uma saída seria usar projetos reais, mas aí perderíamos em outros  
  aspectos. No geral, experimentos em ES são pouco generalizáveis, então  
  pode ser que tenhamos que aceitar e conviver com isso.
* > Que eu tenha conhecimento, não existem ferramentas que comparem diretórios e  
  > arquivos ao mesmo tempo, Não permitindo comparação estatísticas com algumas  
  > determinadas situações, como por exemplo, movimentação de métodos.

### AMEAÇA A VALIDADE EXTERNA

* A ausência de experimentos utilizando grandes projetos deixa a dúvida se o resultado também será satisfatório nestes cenários. Gerando limitação na generalização do resultado.

 Uma saída seria usar projetos reais, mas aí perderíamos em outros  
aspectos. No geral, experimentos em ES são pouco generalizáveis, então  
pode ser que tenhamos que aceitar e conviver com isso.

Que eu tenha conhecimento, não existem ferramentas que comparem diretórios e  
arquivos ao mesmo tempo, Não permitindo comparação estatísticas com algumas  
determinadas situações, como por exemplo, movimentação de métodos.

* No mesmo contexto, o tipo de experimento e a escolha da métrica a ser utilizada pode ser uma ameaça a validade, mas não vejo como evitar, uma vez que algum tipo de experimento deve ser escolhido.
* Não é possível representar todas as situações possíveis de manutenção e evolução de software. Teremos, a principio, a base do livro do Fowler.

### AMEAÇA A VALIDADE CONCLUSÃO

* Não existem ferramentas que comparem diretórios e arquivos ao mesmo tempo, Não permitindo comparação estatísticas em algumas situações, como por exemplo, movimentação de métodos.
* Tamanho limitado da amostra.

### AMEAÇA A VALIDADE DE CONSTRUÇÃO

* Todos testes serão realizados com a configuração default do sistema.

 Tente pensar um pouco nessas colocações e escrever uma versão  
inicial do planejamento do experimento, onde vc estabelece como ele  
deve ser executado e o que deve ser medido em cada momento. Dê uma  
olhada novamente na dissertação do Marcelo para ver como ele fez isso  
(só como exemplo do estilo). Ah, e lembre que escrever bem isso não é  
tempo jogado fora, pois entra na dissertação. Então assim que tiver  
isso me manda que aí refinamos sobre esse plano para batermos o  
martelo