INTERACTION NETWORK DYNAMICS AND INFORMATION TRANSMISSION IN DROSOPHILA

Pasquaretta C, Klenschi E, Pansanel J, Battesti M, Mery F and Sueur C

Journées SUCCES

Rencontres
Scientifiques des Utilisateurs de Calcul intensif, de Cloud Et de Stockage

Why the GRID

Biology

- Some experiments produce large datasets – i.e. DATA MINING
- Agent-based modeling approach
- Simulations and mathematical models

Why the GRID

Biology

- Some experiments produce large datasets – i.e. DATA MINING

New powerful technologies

- Agent-based modeling approach
- Simulations and mathematical models

Why the GRID

Biology

- Some experiments produce large datasets – i.e. DATA MINING

- Agent-based modeling approach

- Simulations and mathematical models

Why the GRID

Table 1
Average and relative run time (Seconds).

	Mac			Windows		
Language	Version/Compiler	Time	Rel. Time	Version/Compiler	Time	Rel. Time
Cþ þ	GCC-4.9.0	0.73	1.00	Visual Cþ þ 2010	0.76	1.00
	Intel Cþ þ 14.0.3	1.00	1.38	Intel Cþ þ 14.0.2	0.90	1.19
	Clang 5.1	1.00	1.38	GCC-4.8.2	1.73	2.29
Fortran	GCC-4.9.0	0.76	1.05	GCC-4.8.1	1.73	2.29
	Intel Fortran 14.0.3	0.95	1.30	Intel Fortran 14.0.2	0.81	1.07
Java	JDK8u5	1.95	2.69	JDK8u5	1.59	2.10
Julia	0.3.7	1.91	2.62	0.3.7	1.80	2.37
Matlab	2014a	7.91	10.88	2014a	6.74	8.92
Python	Pypy 2.2.1	31.90	43.86	Pypy 2.2.1	34.14	45.16
	CPython 2.7.6	195.87	269.31	CPython 2.7.4	117.40	155.31
R	3.1.1, compiled	204.34	280.90	3.1.1, compiled	184.16	243.63
	3.1.1, script	345.55	475.10	3.1.1, script	371.40	491.33
Mathematica	9.0, base	588.57	809.22	9.0, base	473.34	626.19
Matlab, Mex	2014a	1.19	1.64	2014a	0.98	1.29
Rcpp	3.1.1	2.66	3.66	3.1.1	4.09	5.41
Python	Numba 0.13	1.18	1.62	Numba 0.13	1.19	1.57
	Cython	1.03	1.41	Cython	1.88	2.49
Mathematica	9.0, idiomatic	1.67	2.29	9.0, idiomatic	2.22	2.93

Borağan Aruoba and Fernández-Villaverde 2015
J of Economic Dynamics & Control

Access to the GRID

1) DATA mining

```
calcoladistanza <- function( ID, X, Y, Time ) {
  if ( !( length(ID) == length(X) & length(X) == length(Y) & length(Y) == length(Time) ) )
stop("Le variabili di input non hanno la stessa lunghezza.")
  ID <- as.factor(ID)
  Time <- as.factor(Time)
  dist <- array( dim = c( length(levels(ID)), length(levels(ID)), length(levels(Time)) ) )
  dimnames(dist) <- list( levels(ID), levels(ID), levels(Time) )
  for (t in 1:dim(dist)[3])
    for (j in 1:dim(dist)[2])
      for (i in 1:dim(dist)[1])
        if (j<i)  dist[i,j,t] <- sqrt( (X[dim(dist)[3]*(i-1)+t]-X[dim(dist)[3]*(j-1)+t])^2 +
(Y[dim(dist)[3]*(i-1)+t]-Y[dim(dist)[3]*(j-1)+t])^2 )

## output
  dist
}
```

2) Dynamic agent-based model of interactions

```
groupOptions <-sienaAlgorithmCreate(useStdInits=FALSE,
dolby=FALSE,n3=4000,cond=FALSE,firstg=0.01,projname='T-NO',seed=1234567)

groupResults <- siena07(groupOptions,data=groupData,effects=timeheterogenousmodel,
       batch=TRUE,verbose=FALSE,useCluster=TRUE,initC=TRUE,nbrNodes=8,returnDeps=TRUE)
```

Access to the GRID

1) DATA mining

```
calcoladistanza <- function( ID, X, Y, Time ) {
  if ( !( length(ID) == length(X) & length(X) == length(Y) & length(Y) == length(Time) ) )
stop("Le variabili di input non hanno la stessa lunghezza.")
  ID <- as.factor(ID)
  Time <- as.factor(Time)
  dist <- array( dim = c( length(levels(ID)), length(levels(ID)), length(levels(Time)) ) )
  dimnames(dist) <- list( levels(ID), levels(ID), levels(Time) )
  for (t in 1:dim(dist)[3])
    for (j in 1:dim(dist)[2])
      for (i in 1:dim(dist)[1])
        if (j<i)  dist[i,j,t] <- sqrt( (X[dim(dist)[3]*(i-1)+t]-X[dim(dist)[3]*(j-1)+t])^2 +
(Y[dim(dist)[3]*(i-1)+t]-Y[dim(dist)[3]*(j-1)+t])^2 )

## output
  dist
}
```

2) Dynamic agent-based model of interactions

```
groupOptions <-sienaAlgorithmCreate(useStdInits=FALSE,
dolby=FALSE,n3=4000,cond=FALSE,firstg=0.01,projname='T_NO',seed=1234567)

groupResults <- [illegible],
      batch=TRUE,verbose=FALSE,useCluster=TRUE,initC=TRUE,nbrNodes=8,returnDeps=TRUE)
```

RSiena about 300 hours running

Access to the GRID

setwd("/opt/sbg/grand-est/ui9_data2/**cpasquar**/T-NO")

Jérome Pansanel

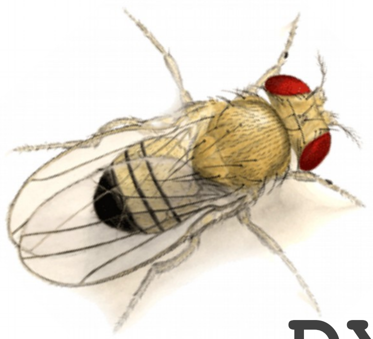
Parallel computing

2) Dynamic agent-based model of interactions

```
groupOptions <-sienaAlgorithmCreate(useStdInits=FALSE,
dolby=FALSE,n3=4000,cond=FALSE,firstg=0.01,projname='T_NO',seed=1234567)

groupResults <- siena0                                    usmodel,
      batch=TRUE,verbose=FALSE,useCluster=TRUE,initC=TRUE,nbrNodes=8,returnDeps=TRUE)
```

About 300 hours running

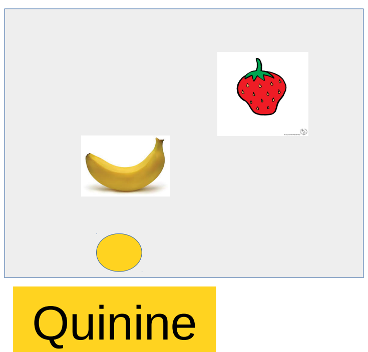


INTERACTION NETWORK DYNAMICS AND INFORMATION TRANSMISSION IN DROSOPHILA

Experiment

Phase I
Training (8 h)

8 Demonstrators



Phase II
Social transmission (4 h)

Phase III
Test (4 h)

Battesti *et al.* 2012 *Curr Biol*

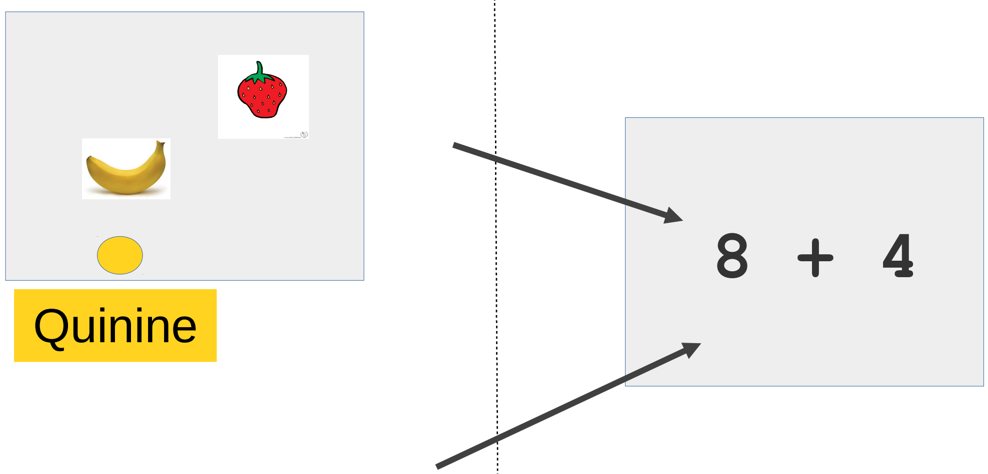
Experiment

Phase I
Training (8 h)

Phase II
Social transmission (4 h)

Phase III
Test (4 h)

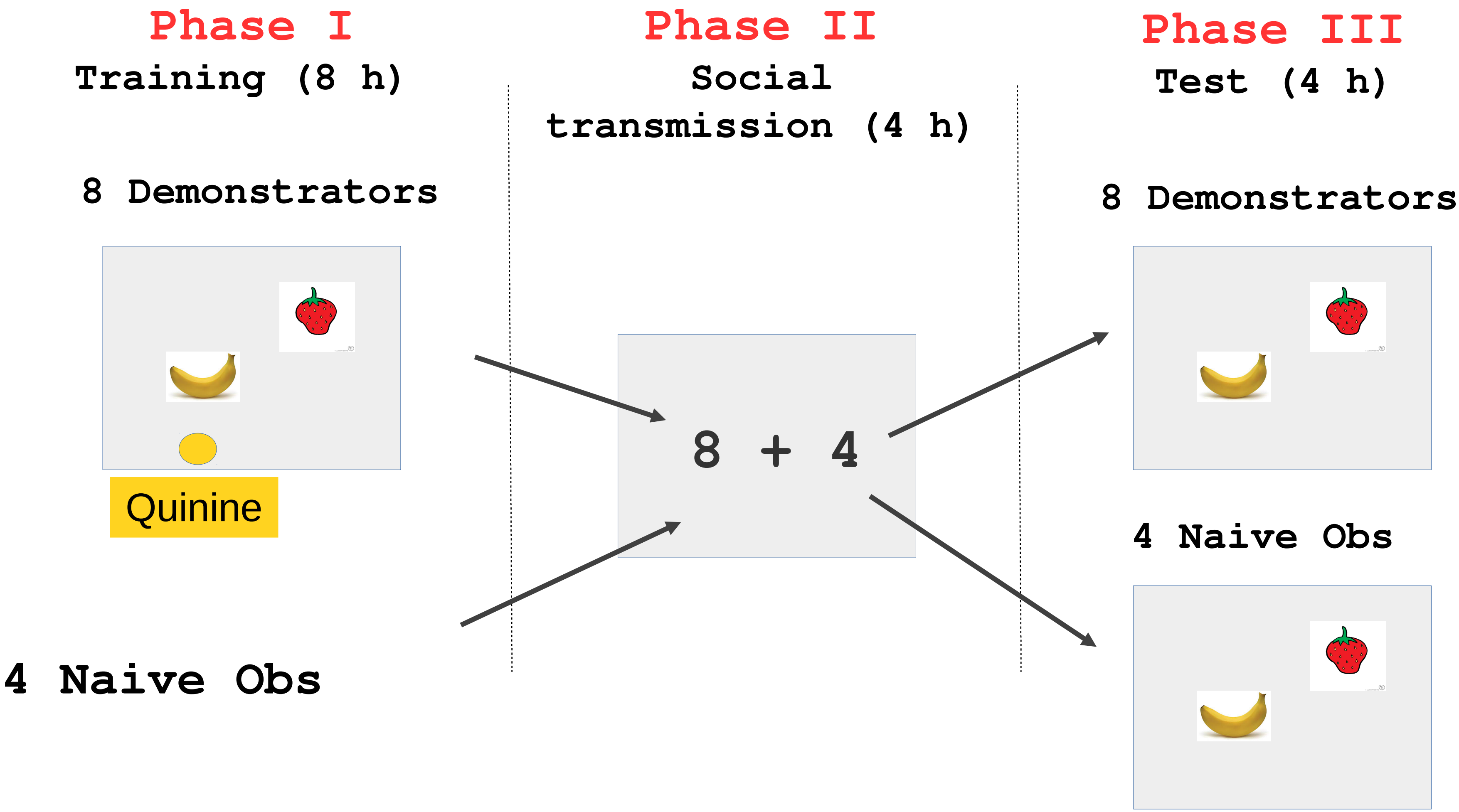
8 Demonstrators



4 Naive Obs

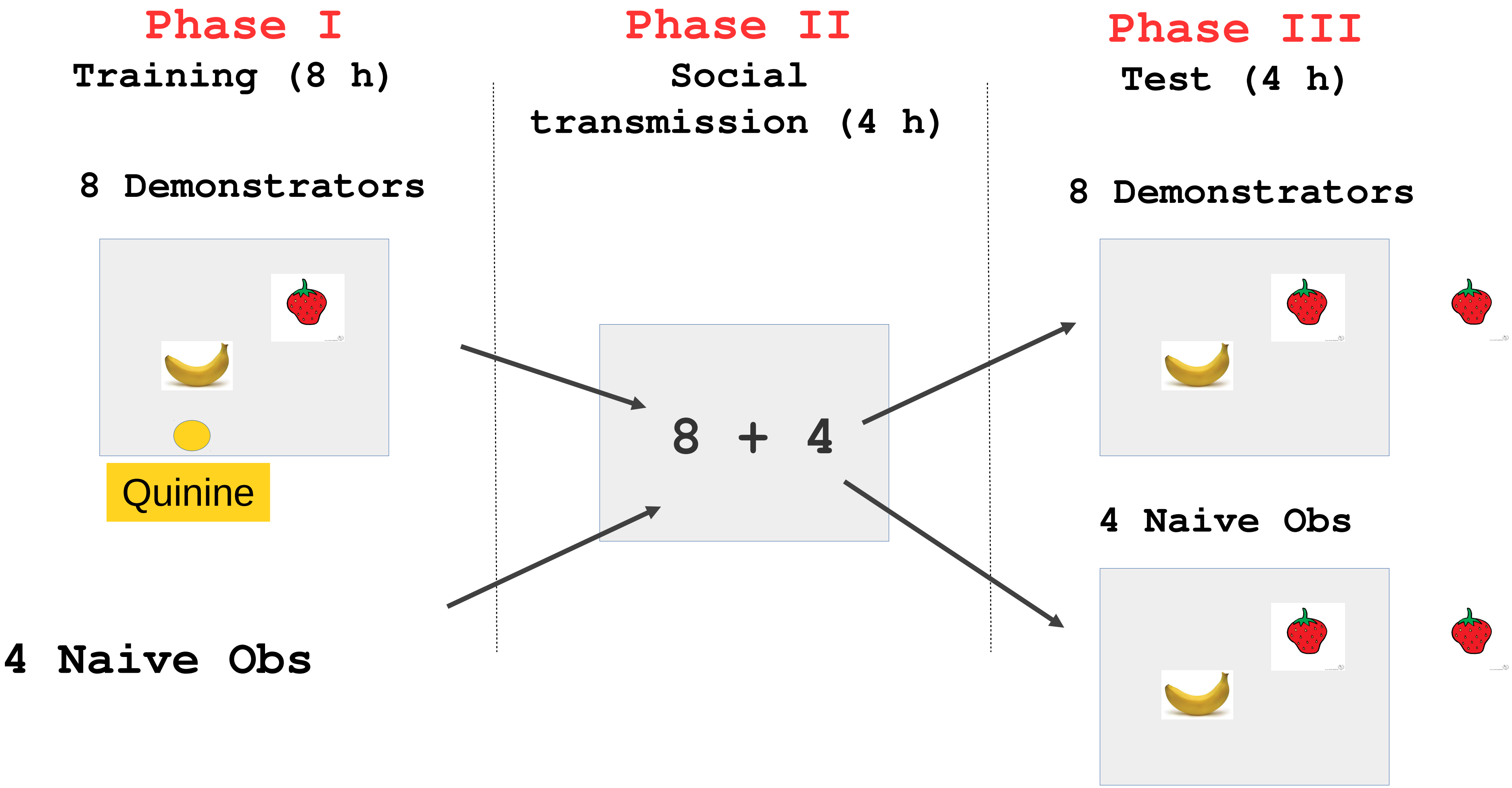
Battesti *et al.* 2012 *Curr Biol*

Experiment



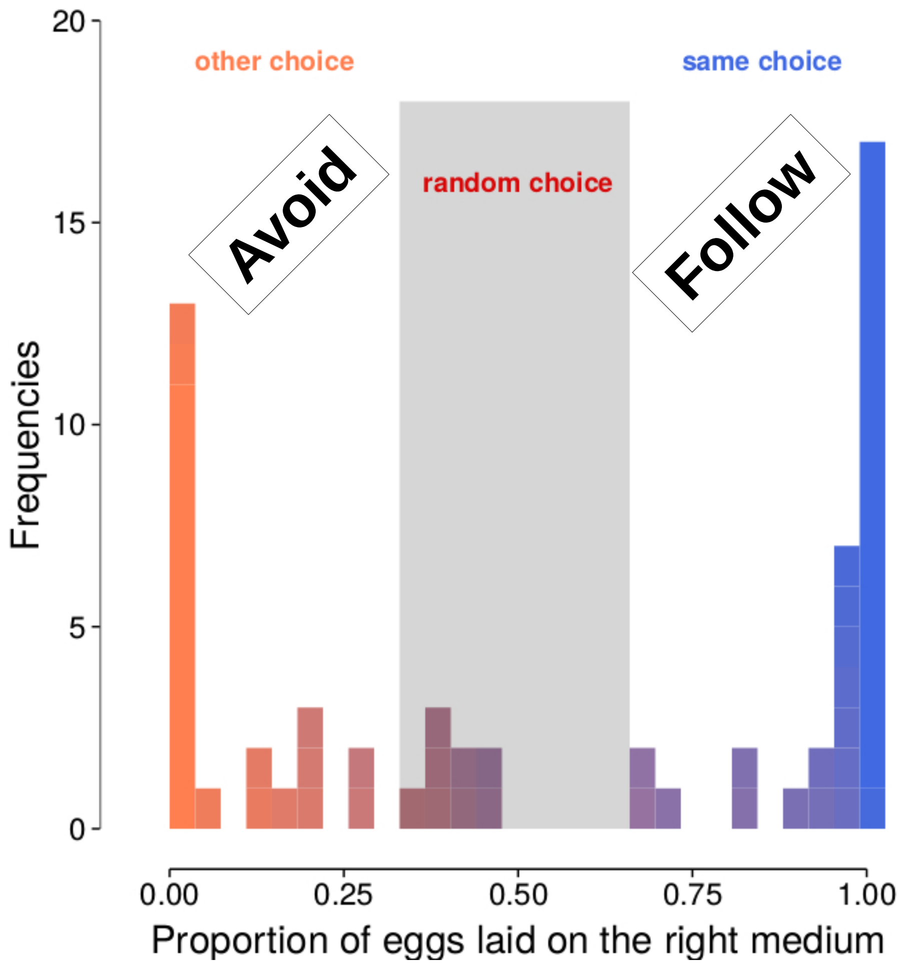
Battesti *et al.* 2012 *Curr Biol*

Experiment



Battesti *et al.* 2012 *Curr Biol*

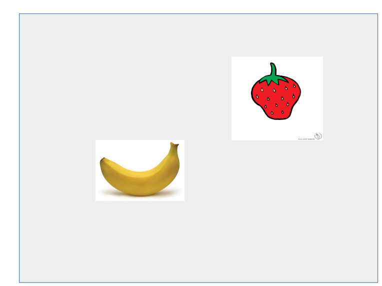
Experiment



Phase III

Test (4 h)

4 Naive Obs



Experiment

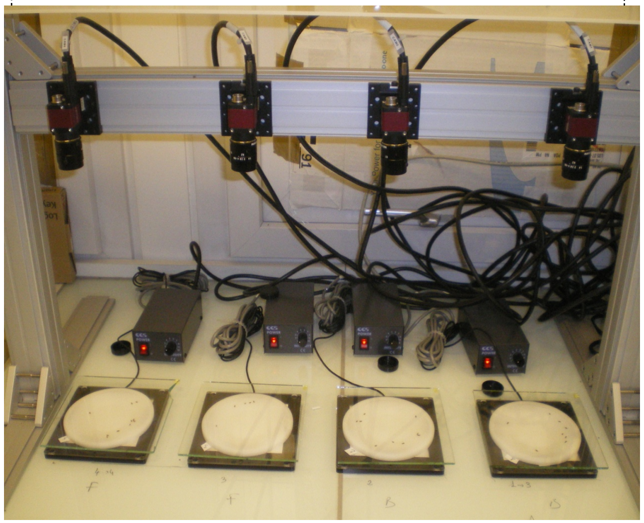
Phase I
Training (8 h)

8 Informed

4 Uninformed flies

Phase II
Social transmission (4 h)

8 + 4



Phase III
Test (4 h)

8 Informed

4 Uninformed

Experiment

Phase I

Phase II

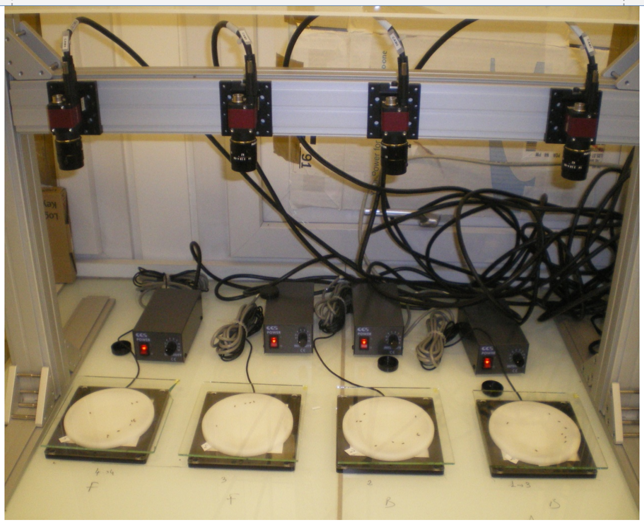
Phase III

Do contact interactions explain social learning?

8 + 4

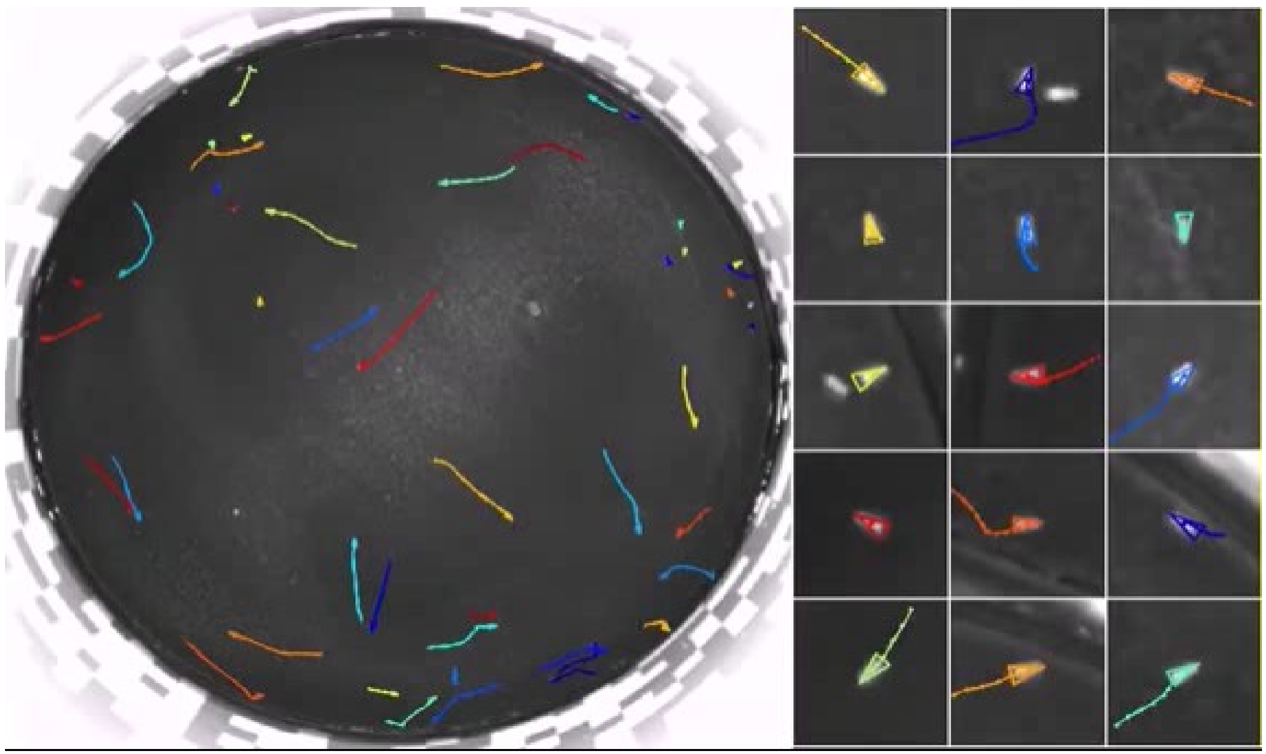
4 Uninformed

4 Uninformed flies



Why the GRID

1) DATA mining



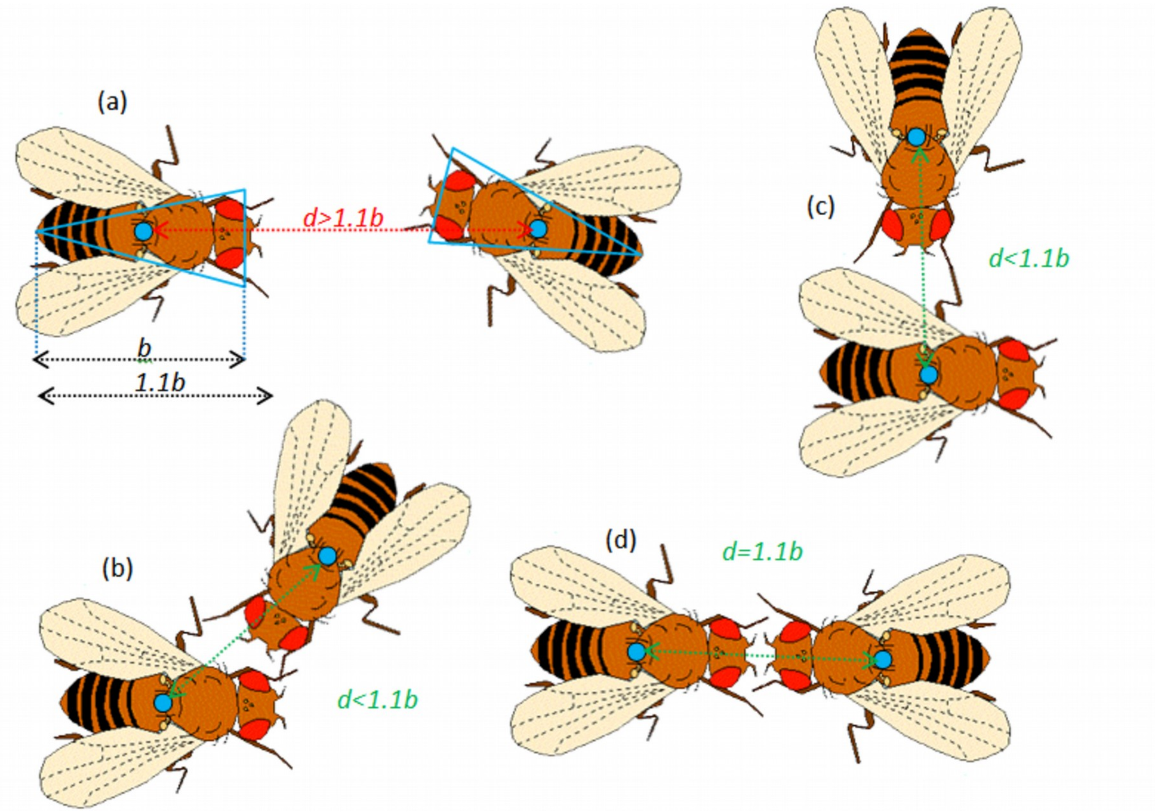
- **10** spatial positions for each fly **per second**
- **12 flies**
- **4 hours** video tracking

= 1 728 000

*** 62 videos**

Why the GRID

1) DATA mining



i) **Spatial constrain:**
Proximity (*d*) < 1.1 average body lengths (*b*)

ii) **Temporal constrain:**
Contact duration at least 0.5 s

iii) **Speed rules:**
The initiator was the faster fly over the 0.5 s that preceded each interaction

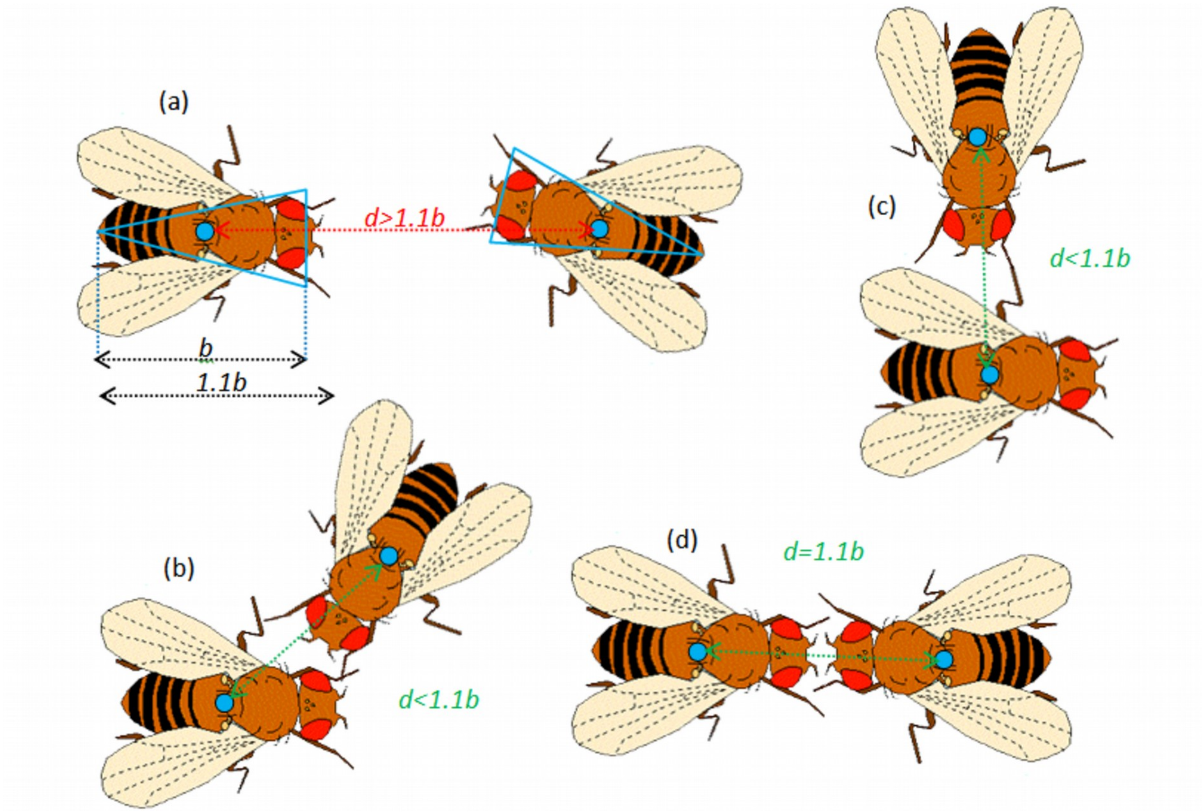
Why the GRID

1) DATA mining

Receivers

Senders

	A	B	C	D
A	–	3	0	15
B	2	–	0	0
C	0	0	–	1
D	1	0	7	–



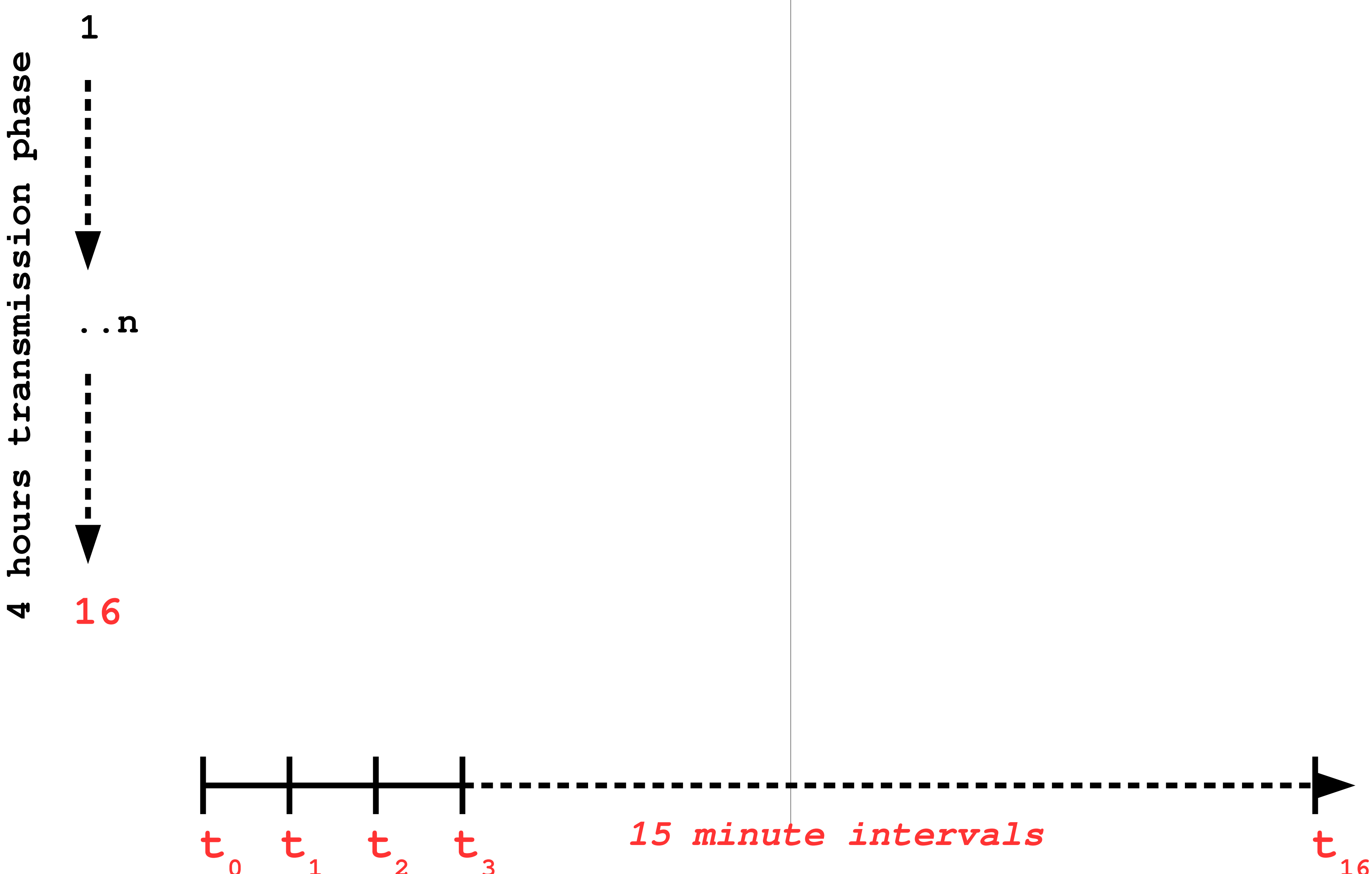
i) **Spatial constrain:**
Proximity (*d*) < 1.1 average body lengths (*b*)

ii) **Temporal constrain:**
Contact duration at least 0.5 s

iii) **Speed rules:**
The initiator was the faster fly over the 0.5 s that preceded each interaction

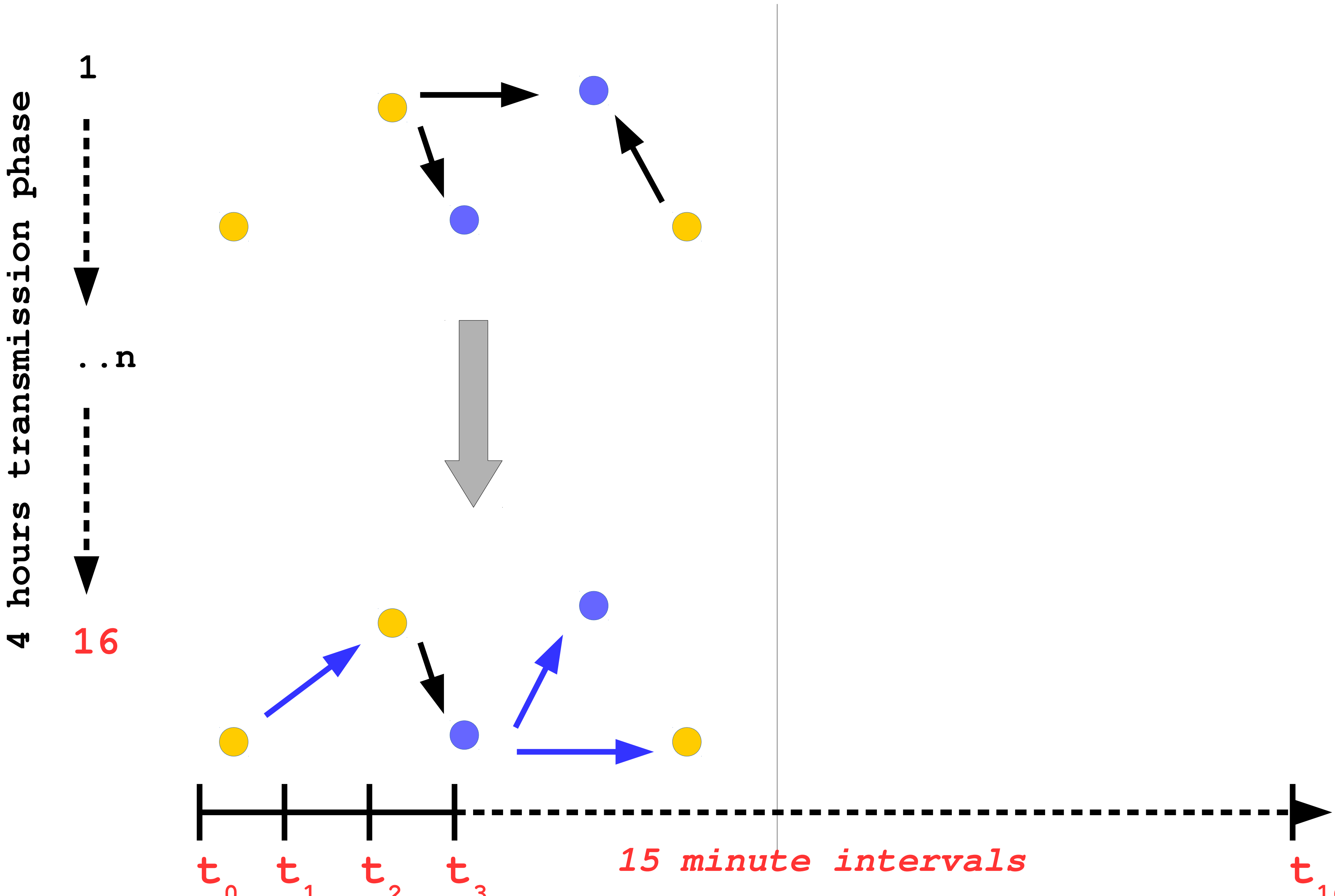
Why the GRID

2) Dynamic agent-based model (RSiena package)



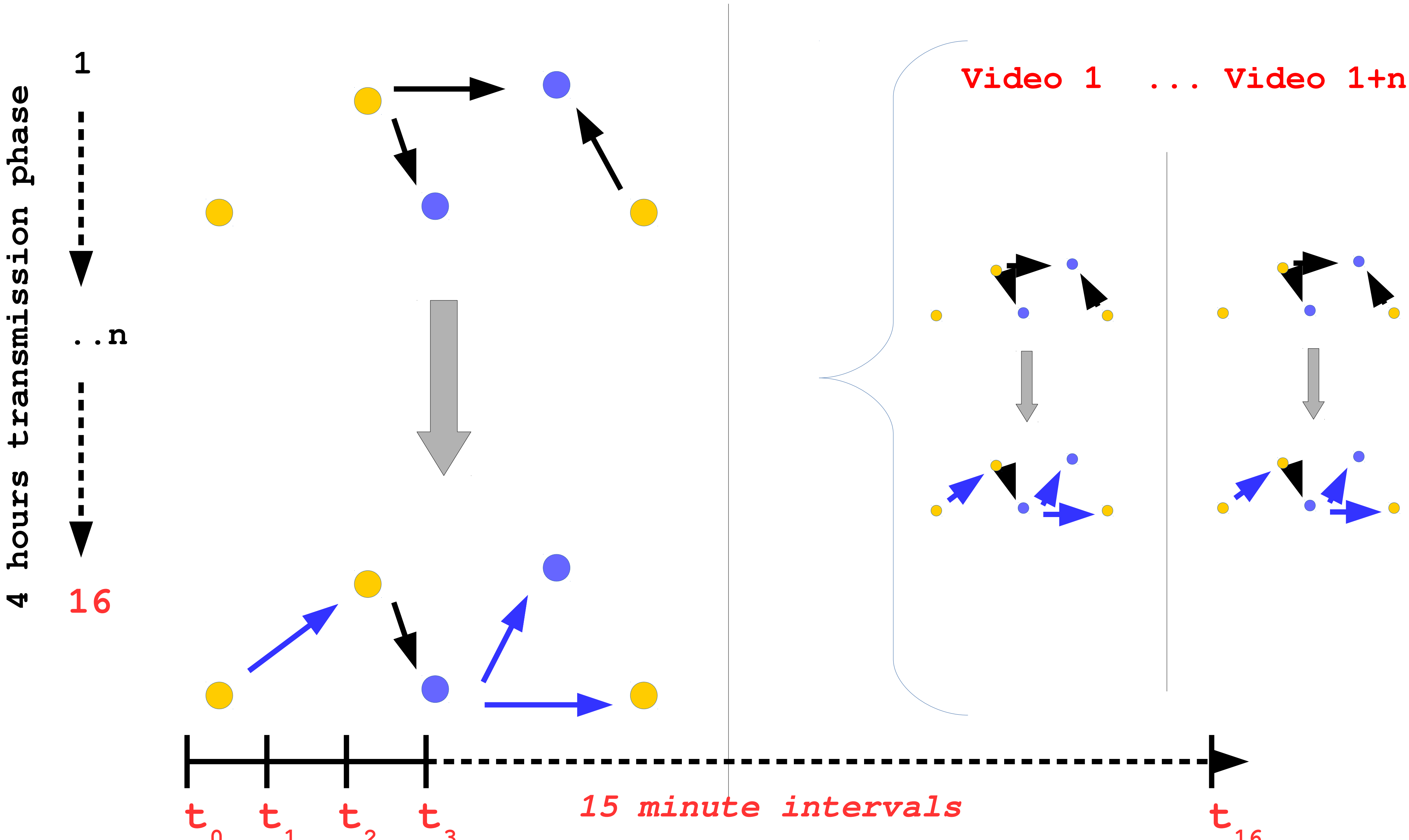
Why the GRID

2) Dynamic agent-based model (RSiena package)



Why the GRID

2) Dynamic agent-based model (RSiena software)

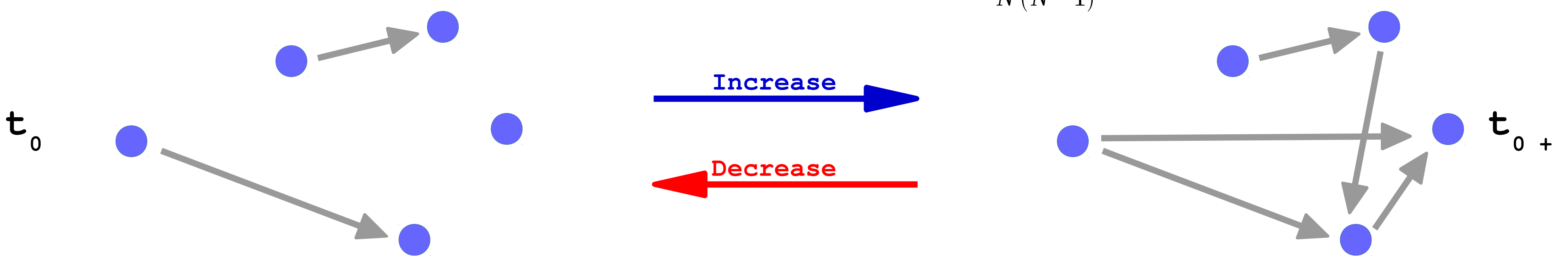


Testing dynamic effects

Temporal Dynamic

1) **Density** used as covariate

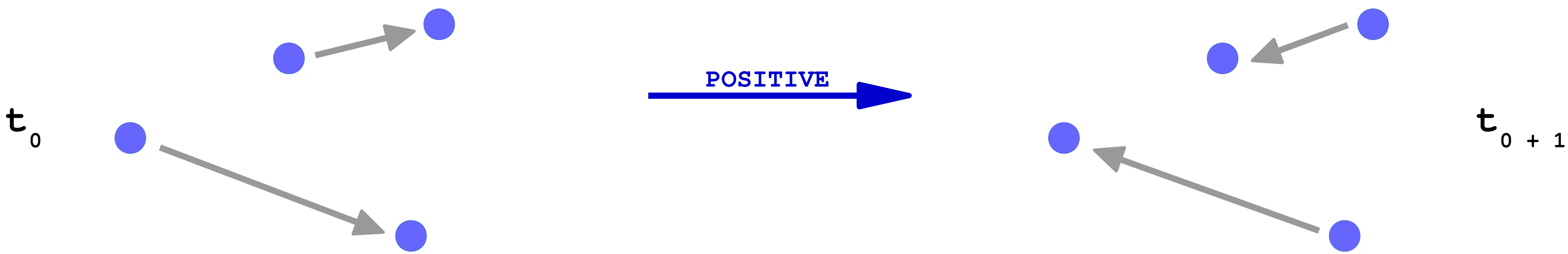
$$D = \frac{links}{N(N-1)}$$



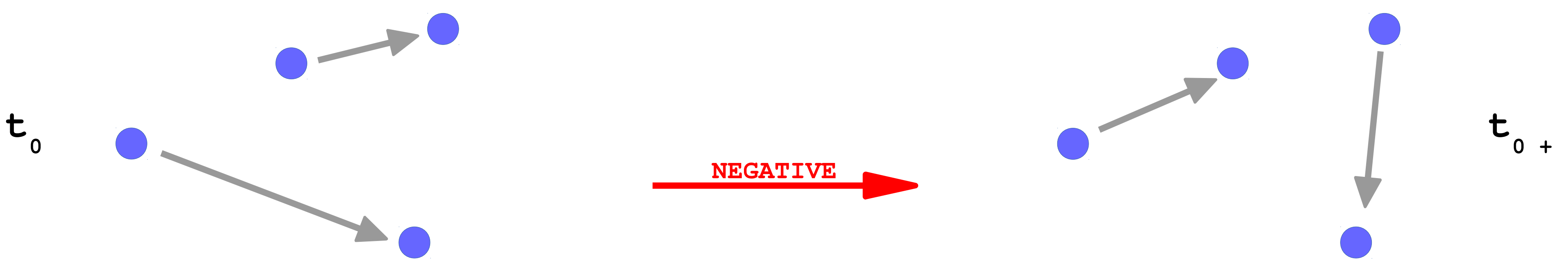
The realized number of links divided by the number of maximum possible links in the network

Temporal Dynamic

2) Reciprocity

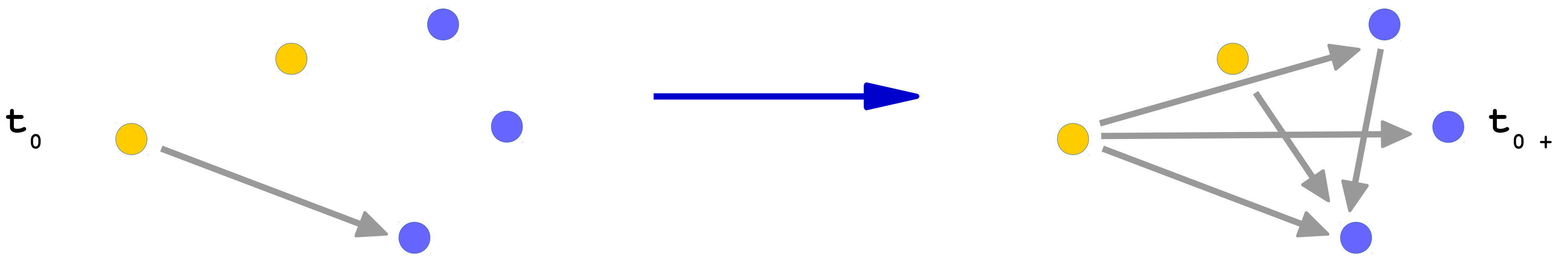


defined by the number of reciprocated interactions



Temporal Dynamic

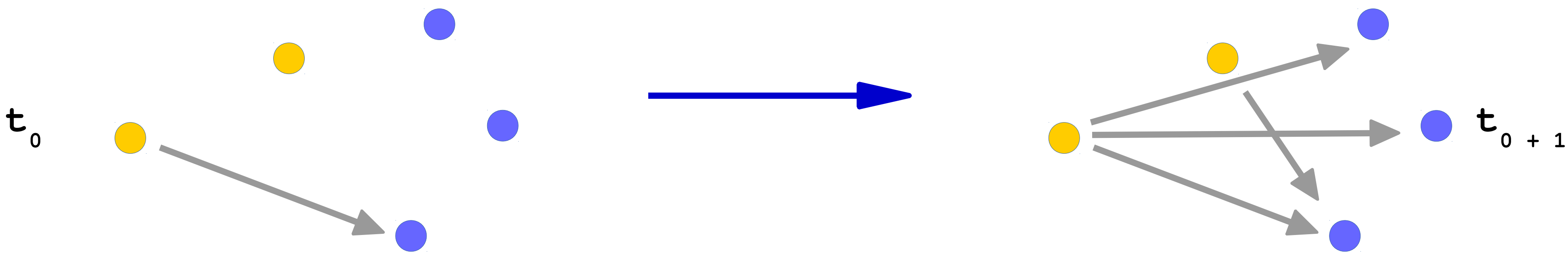
3) covariate interactions **received** (Alter)



It expresses which class of actors **receives** more interactions

Temporal Dynamic

4) covariate interactions **started** (Ego)



It expresses which class of actors **starts** more interactions

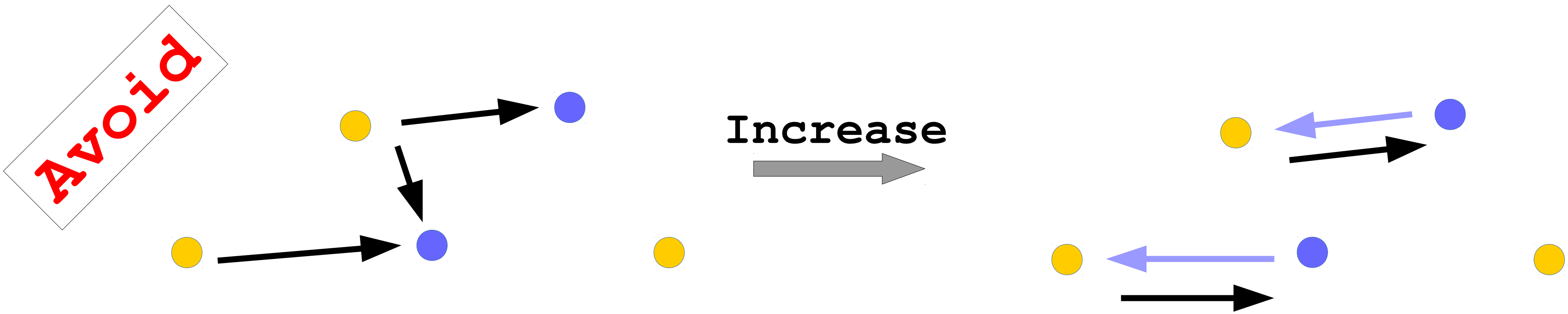
Results

1) Density

Naive Obs   Demonstrators

Results

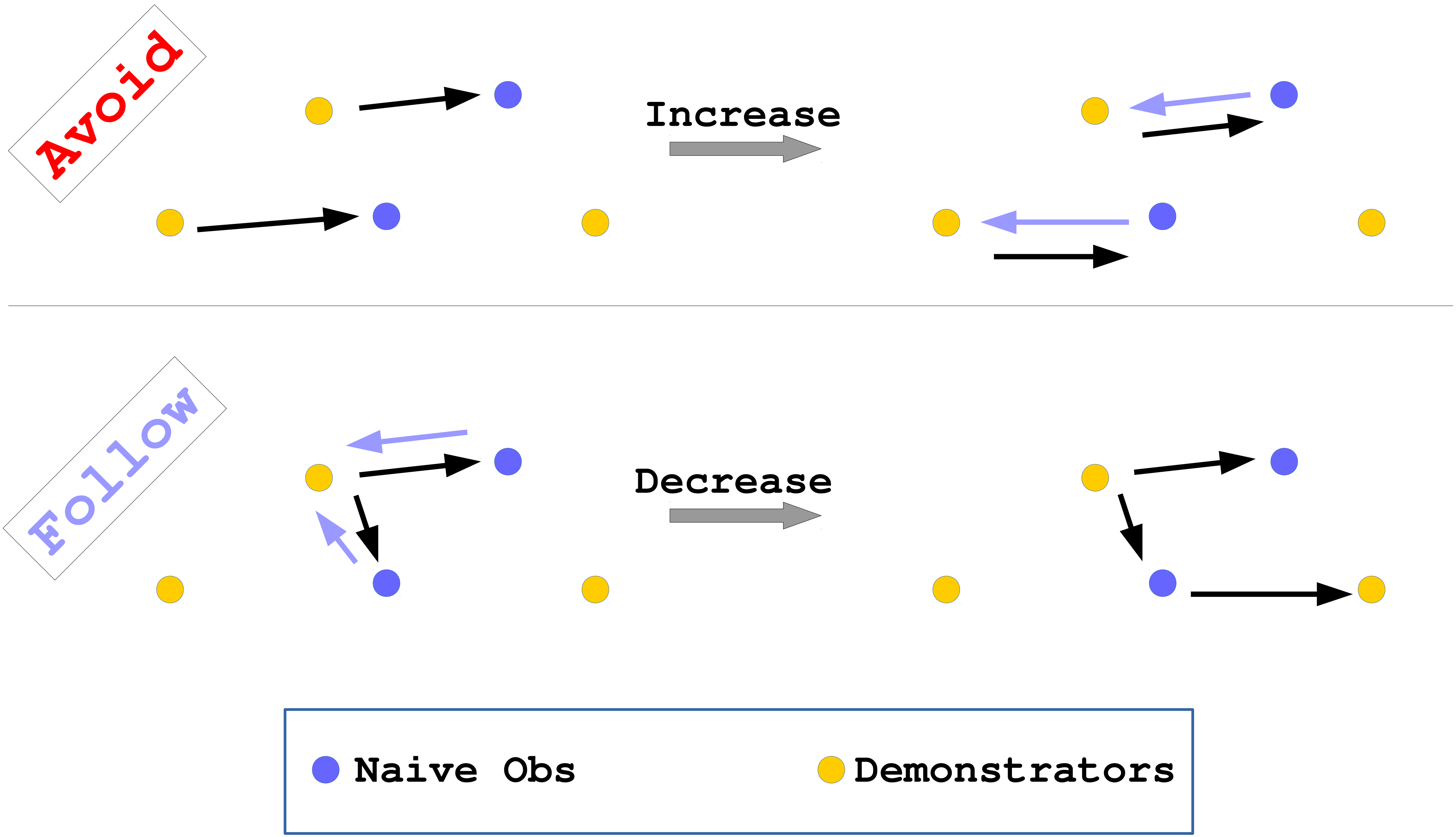
2) Reciprocity



Naive Obs   Demonstrators

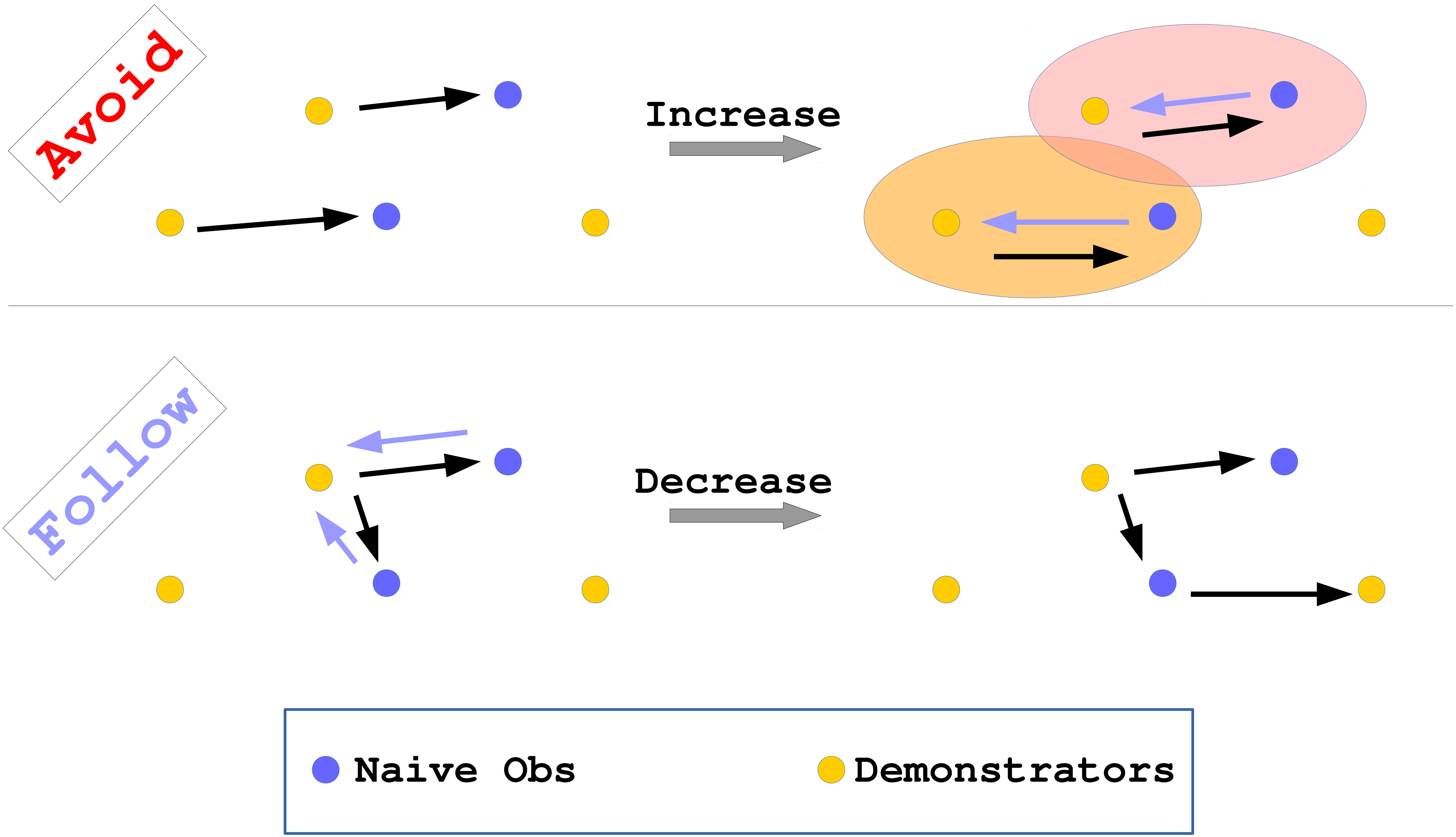
Results

2) Reciprocity

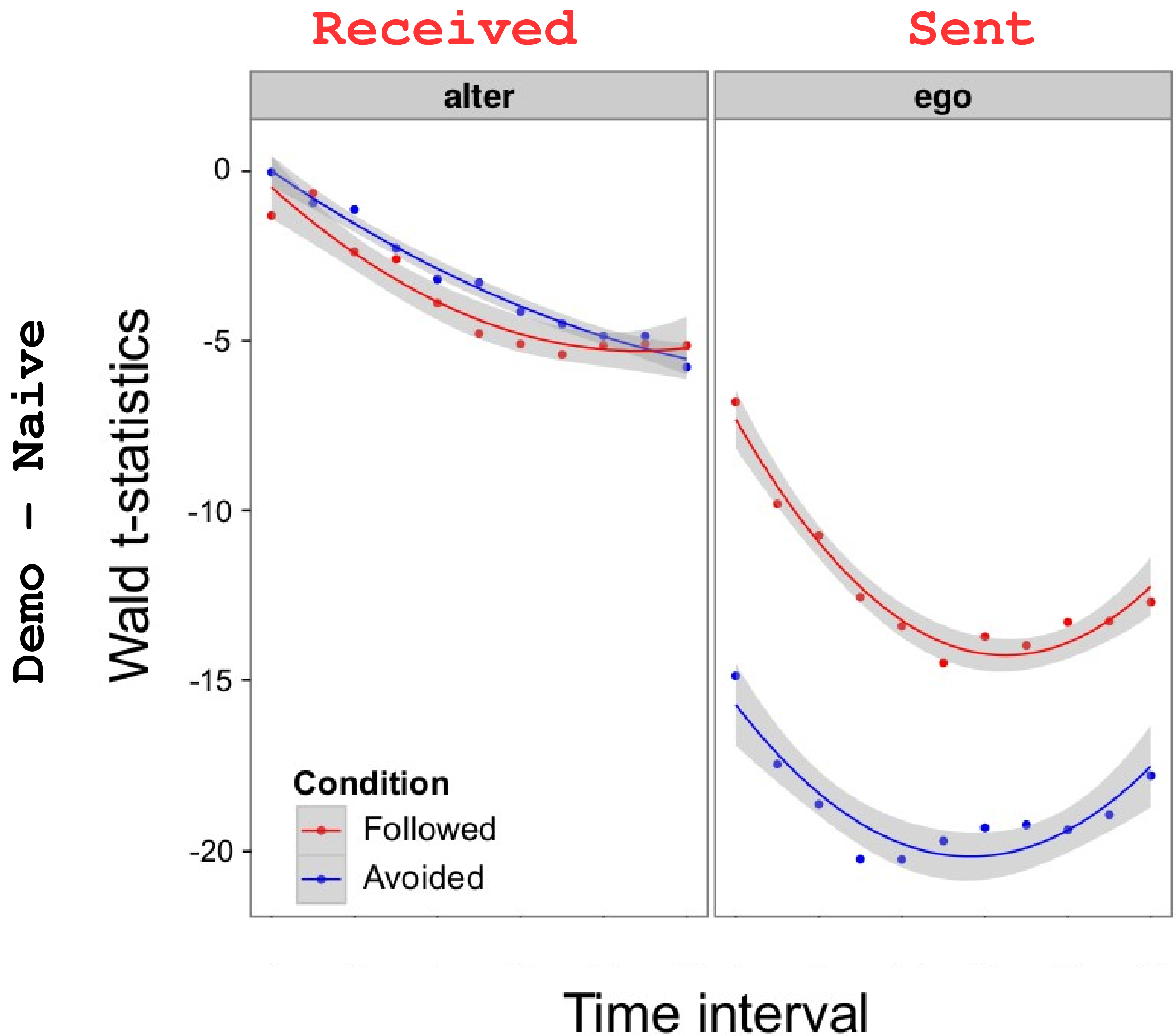


Results

2) Reciprocity

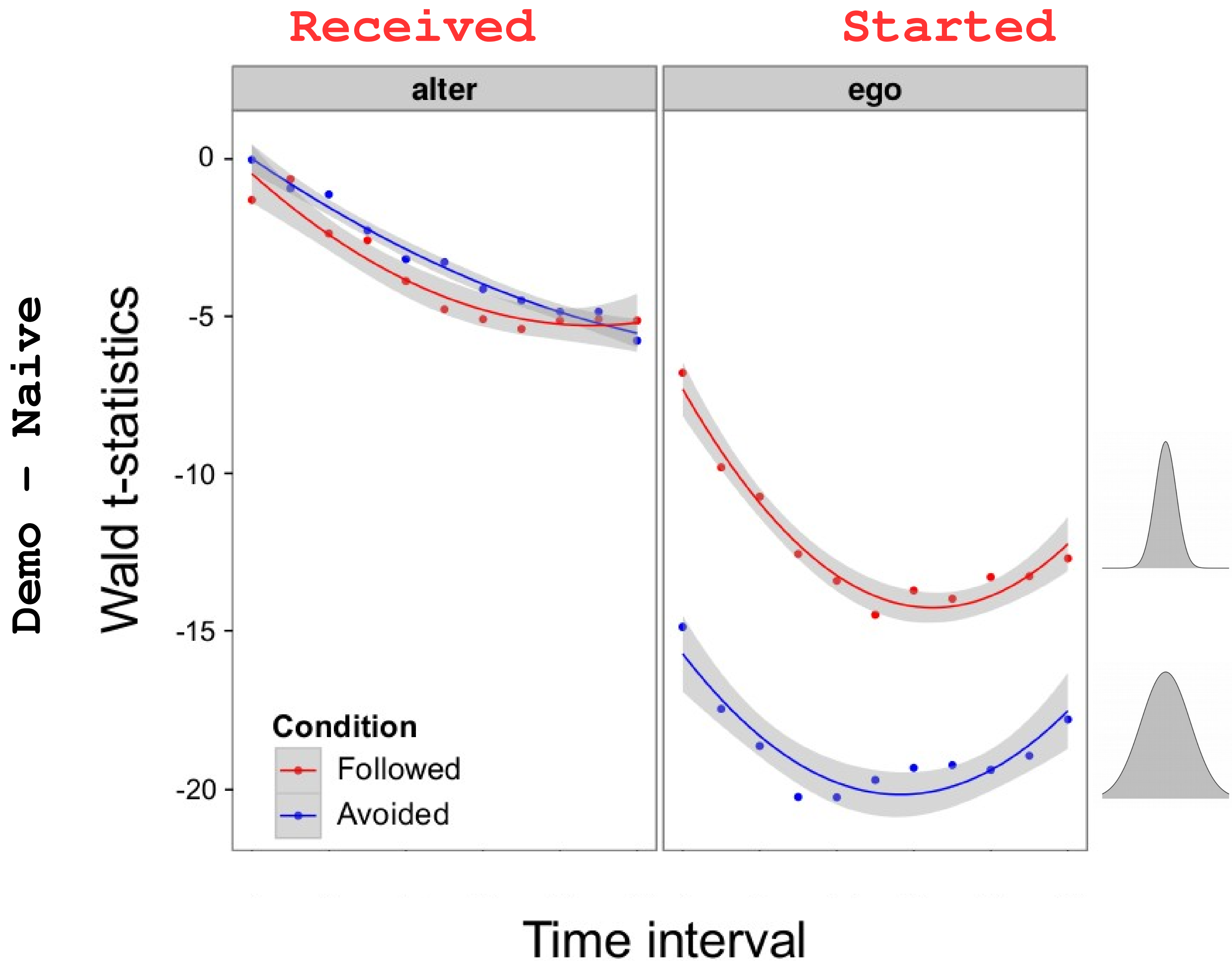


Results



Received
Sent
alter
ego
Demo - Naive
Wald t-statistics
0
-5
-10
-15
-20
Condition
Followed
Avoided
Time interval

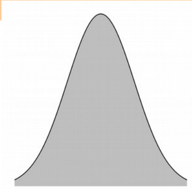
Results

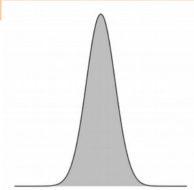


Received
Started
alter
ego
Demo – Naive
Wald t-statistics
0
-5
-10
-15
-20
Condition
Followed
Avoided
Time interval

Sum - up

	Avoid	Follow
Reciprocity	Increase	Decrease
Delta Interactions received	Naive more than Demo	Naive more than Demo
Delta Interactions started	Naive more than Demo	Naive more than Demo





- **Reciprocity** decrease probably brings to less clustered networks
- **Reduced variance** between Demo and Naive implies an increase in the homogeneity of behaviours

Publication
April 2016

frontiers in Psychology

ORIGINAL RESEARCH
published: 19 April 2016
doi: 10.3389/fpsyg.2016.00539

CrossMark

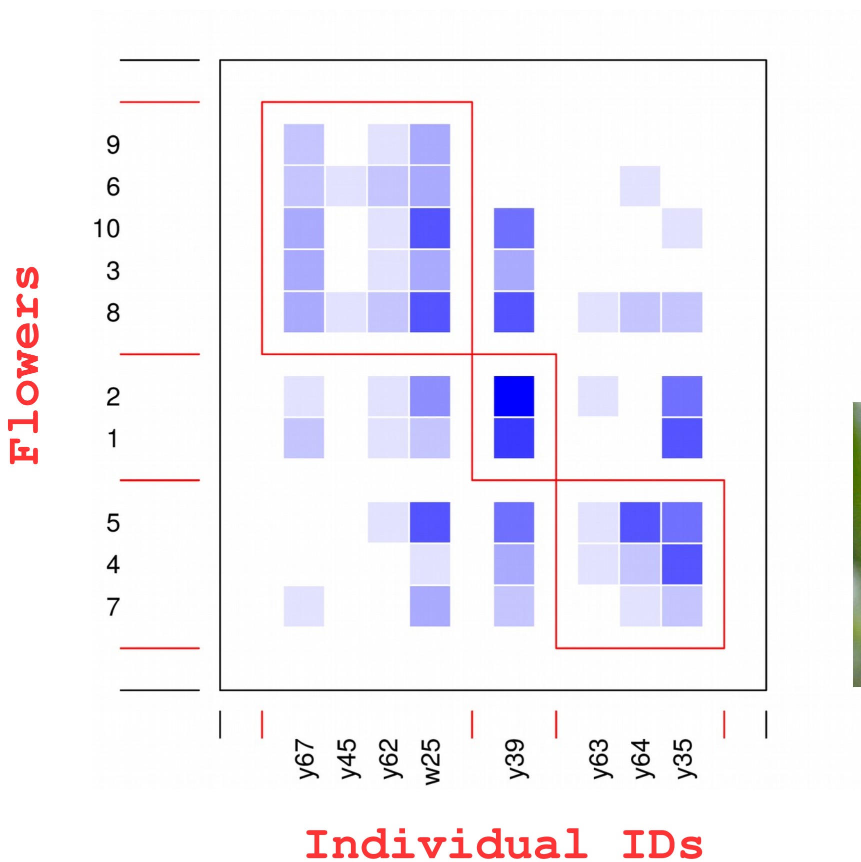
Understanding Dynamics of Information Transmission in Drosophila melanogaster Using a Statistical Modeling Framework for Longitudinal Network Data (the RSiena Package)

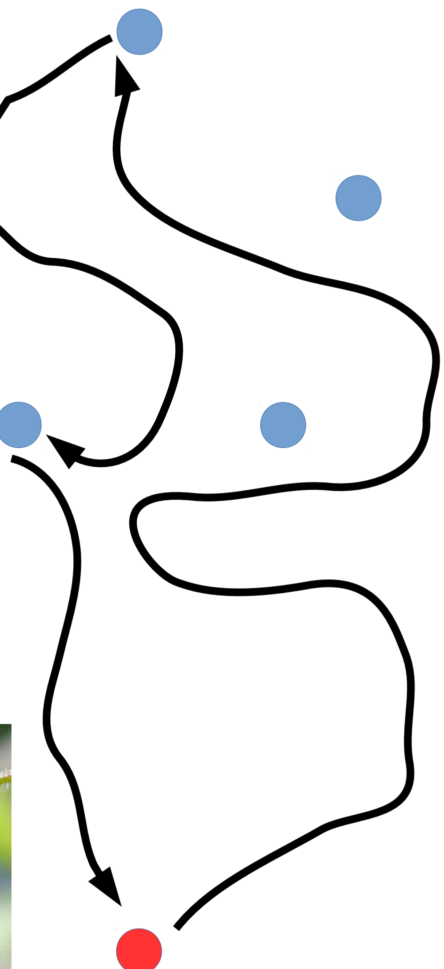
Cristian Pasquaretta[1,2*†], Elizabeth Klenschi[1,2†], Jérôme Pansanel[1,2], Marine Battesti[3], Frederic Mery[3] and Cédric Sueur[1,2]

Today with the GRID

Resource partitioning in pollinators

Dr. Mathieu Lihoreau (CNRS-CRCA-Toulouse)
Dr. Raphael Jeanson (CNRS-CRCA-Toulouse)
Prof. Lars Chittka (Queen Mary, University of London)





Today with the GRID

Jérome Pansanel

```
mat=NULL
Nsimul=100

for (Nflowers in c(5,10,100,500))
{

  Seq<-function (meanAct,Ps)
  {
    vec<-rzipois(10, lambda = meanAct,Ps)
    vec<-(round((vec/length(which(vec>0)))))
    vec[which(vec=="NaN")]<-0
    vec<-as.numeric(vec)

    vec2=NULL
    if(Nflowers==5)
    {
      vec2=c(sum(vec[1:2]),sum(vec[3:4]),sum(vec[5:6]),sum(vec[7:8]),sum(vec[9:10]))
      vec=vec2
    }

    if (Nflowers>=100)
      vec=sample(c(vec,rep(0,Nflowers-10)))


    return(vec)
  }

  for (Nind in c(5:10,20,50))
  {
    for (proba_Spec in seq(0,0.9,0.1))
    {
      for (simulations in 1:Nsimul)
      {

        for (N_specialist in 0:Nind)
        {
          N_generalist<-Nind-N_specialist
          if(N_specialist>0 & N_generalist>0)
          { ………………
```

Parallel computing
100 simulations

few hours per simulation

DIRAC
THE INTERWARE

MERCI de votre attention

Funding:

ANR PROJECT DROSONET PI: Dr. Frederic Mery
University of Strasbourg Co-PI: Ass. Prof. Cédric Sueur

GDR RSD CNRS
FRANCE GRILLES
Groupe CALCUL
Grid'5000

Journées SUCCES

Rencontres
Scientifiques des Utilisateurs de Calcul intensif, de Cloud Et de Stockage

Agence Nationale de la Recherche ANR

cnrs dépasser les frontières

UNIVERSITÉ DE STRASBOURG

USIAS