□

APPENDIX

A RJAGS CODE

A.1 Model for section 3

```
model{
## x are the count in each category (11,10,01,00), N is total number
## method 1
    x[1:4] ~ dmulti(p[1:4],N)
    p[1]<-p.111
    p[2]<-p.1dd-p.111
    p[3]<-p.dd1-p.111
    p[4]<-1-p.1dd-p.dd1+p.111
# prior dist'n
    p.1dd~dbeta(1,1)
    p.dd1~dbeta(1,1)
    a1<-2*(max(0,p.1dd+p.dd1-1)-p.1dd*p.dd1)/(p.1dd+p.dd1-2*p.1dd*p.dd1) # range of kappa1
    b1<-2*(min(p.1dd,p.dd1)-p.1dd*p.dd1)/(p.1dd+p.dd1-2*p.1dd*p.dd1) # range of kappa1
    kap1~dunif(a1,b1)
    p.111<-kap1*(p.1dd+p.dd1-2*p.1dd*p.dd1)/2+p.1dd*p.dd1


## method 2
    x[5:8] ~ dmulti(q[1:4],N)
    q[1]<-p.112
    q[2]<-p.2dd-p.112
    q[3]<-p.dd2-p.112
    q[4]<-1-p.2dd-p.dd2+p.112
# prior dist'n, marginals are linked to method 1 by parameter c1
    p.2dd<-p.1dd^c1
    p.dd2<-p.dd1^c1
    c1~dunif(0,10)

    a2<-2*(max(0,p.2dd+p.dd2-1)-p.2dd*p.dd2)/(p.2dd+p.dd2-2*p.2dd*p.dd2)
    b2<-2*(min(p.2dd,p.dd2)-p.2dd*p.dd2)/(p.2dd+p.dd2-2*p.2dd*p.dd2)
    kap2~dunif(a2,b2)
    p.112<-kap2*(p.2dd+p.dd2-2*p.2dd*p.dd2)/2+p.2dd*p.dd2

## method 3
    x[9:12] ~ dmulti(r[1:4],N)
    r[1]<-p.113
    r[2]<-p.3dd-p.113
    r[3]<-p.dd3-p.113
    r[4]<-1-p.3dd-p.dd3+p.113
# prior dist'n, marginals are linked to method 1 by parameter c2
    p.3dd<-p.1dd^c2
    p.dd3<-p.dd1^c2
    c2~dunif(0,10)

    a3<-2*(max(0,p.3dd+p.dd3-1)-p.3dd*p.dd3)/(p.3dd+p.dd3-2*p.3dd*p.dd3)
    b3<-2*(min(p.3dd,p.dd3)-p.3dd*p.dd3)/(p.3dd+p.dd3-2*p.3dd*p.dd3)
    kap3~dunif(a3,b3)
    p.113<-kap3*(p.3dd+p.dd3-2*p.3dd*p.dd3)/2+p.3dd*p.dd3

## compare kappa's for method 1 and 2
    diff12<-kap1-kap2

## compare kappa's for method 1 and 3
    diff13<-kap1-kap3

## compare kappa's for method 2 and 3
    diff23<-kap2-kap3

}
```

A.2 Model for section 4

```
model{
for (i in 1:N1){
# for each record i in 1:N1, x[i,] indicates the outcome category (11,10,01,00)
    x[i,1:4] ~ dmulti(p[i,1:4],1)
    p[i,1]<-(p.1dd[i]+p.dd1[i]+p.ag[i]-1)/2
    p[i,2]<-(p.1dd[i]-p.dd1[i]-p.ag[i]+1)/2
    p[i,3]<-(p.dd1[i]-p.1dd[i]-p.ag[i]+1)/2
    p[i,4]<-(p.ag[i]-p.1dd[i]-p.dd1[i]+1)/2

# marginal models
# gamma[1] is intercept, gamma[2:7] are coefficients for covariates z[i,], a[subj[i]] is random intercept
    logit(p.1dd[i])<-gamma1[1]+inprod(gamma1[2:7],z[i,])+a1[subj[i]]
    logit(p.dd1[i])<-gamma2[1]+inprod(gamma2[2:7],z[i,])+a2[subj[i]]

# model for kappa
# beta[1] is intercept, beta[2:6] are coefficients for rater indicator rater[i,],
# beta[7:8] are coefficients for method indicator meth[i,]
    cag.p[i]<-p.1dd[i]*p.dd1[i]+(1-p.1dd[i])*(1-p.dd1[i])
    p.ag[i]<-cag.p[i]+(1-cag.p[i])*kap[i]
    kap[i]<-beta[1]+inprod(beta[2:6],rater[i,])+inprod(beta[7:8],meth[i,])
}

# prior dist'n for the subject random effect for each subject i in 1:M
for (i in 1:M) {
    a1[i]~dnorm(0, tau_int) # Random intercepts
    a2[i]~dnorm(0, tau_int)
}
sigma_int~dunif(0, 100)
tau_int <- 1/(sigma_int*sigma_int)

# prior dist'n for gamma's
gamma1[1:7] ~ dmnorm(mu0[1:7],iSigma0[1:7,1:7])
gamma2[1:7] ~ dmnorm(mu0[1:7],iSigma0[1:7,1:7])

for(j in 1:7){
    mu0[j] <- 0
    iSigma0[j,j] <- 1/100
}
for(j in 2:7){
    for(k in 1:(j-1)){
        iSigma0[j,k] <- 0
        iSigma0[k,j] <- 0
    }
}

# prior dist'n for beta's
beta[1:8] ~ dmnorm(mu1[1:8],iSigma1[1:8,1:8])

for(j in 1:8){
    mu1[j] <- 0
    iSigma1[j,j] <- 1/100
}
for(j in 2:8){
    for(k in 1:(j-1)){
        iSigma1[j,k] <- 0
        iSigma1[k,j] <- 0
    }
}}
```

B ADDITIONAL RESULTS

Table S1. The posterior distribution summary of c's for all readers under BMDK

readers	c_1			c_2		
	mean	median	95%BCI*	mean	median	95%BCI*
1	1.185	1.157	0.763, 1.763	1.207	1.179	0.78, 1.796
2	2.296	2.199	1.292, 3.861	2.413	2.31	1.362, 4.042
3	1.945	1.882	1.156, 3.095	1.162	1.12	0.665, 1.898
4	1.098	1.077	0.734, 1.581	0.965	0.947	0.64, 1.397
5	1.349	1.297	0.771, 2.223	1.288	1.24	0.731, 2.124
6	1.178	1.145	0.716, 1.833	1.386	1.346	0.857, 2.143

*BCI: Bayesian Credible Interval

Table S2. The posterior distribution summary of β's, γ's for 2-stage modeling and joint modeling

Model	Parameter		Two-stage			Joint		
			Mean	Median	95%BCI***	Mean	Median	95%BCI***
Round1*	Intercept	γ_{10}	2.826	2.883	1.361, 3.672	2.092	2.128	0.831, 3.191
	BMI	γ_{11}	-0.985	-0.994	-1.779, -0.089	-0.983	-0.951	-1.783, -0.431
	sex	γ_{12}	-1.059	-1.148	-2.124, 0.278	-0.819	-0.874	-2.136, 0.624
	pack_years	γ_{13}	0.001	0.001	-0.036, 0.038	0.002	0.002	-0.022, 0.028
	age	γ_{14}	1.096	1.054	0.332, 1.890	0.775	0.758	0.150, 1.431
	HC vs SC	γ_{15}	0.701	0.706	0.331, 1.090	0.679	0.669	0.319, 1.066
	FSR vs SC	γ_{16}	0.080	0.076	-0.308, 0.489	-0.039	-0.040	-0.384, 0.290
Round2*	Intercept	γ_{20}	2.844	2.613	0.849, 5.330	1.837	1.832	1.016, 2.656
	BMI	γ_{21}	-0.916	-0.927	-1.785, -0.126	-0.726	-0.752	-1.256, -0.127
	sex	γ_{22}	-0.739	-0.501	-3.501, 0.946	-0.191	-0.225	-1.151, 0.780
	pack_years	γ_{23}	0.005	0.006	-0.035, 0.042	0.006	0.006	-0.022, 0.034
	age	γ_{24}	0.437	0.402	-0.425, 1.362	0.640	0.646	0.030, 1.145
	HC vs SC	γ_{25}	0.337	0.341	-0.098, 0.768	0.410	0.423	-0.087, 0.809
	FSR vs SC	γ_{26}	-0.529	-0.523	-0.945, -0.122	-0.407	-0.391	-0.815, -0.075
Kappa model**	Intercept	β_0	0.504	0.508	0.300, 0.695	0.159	0.160	0.053, 0.260
	Rater2	β_1	-0.246	-0.246	-0.505, 0.009	-0.017	-0.011	-0.148, 0.087
	Rater3	β_2	-0.165	-0.164	-0.413, 0.076	-0.027	-0.022	-0.153, 0.083
	Rater4	β_3	-0.158	-0.156	-0.400, 0.082	0.013	0.009	-0.062, 0.103
	Rater5	β_4	-0.179	-0.178	-0.430, 0.066	0.022	0.016	-0.042, 0.110
	Rater6	β_5	-0.101	-0.102	-0.342, 0.141	0.004	0.003	-0.091, 0.098
	HC vs SC	β_6	0.101	0.102	-0.080, 0.278	0.061	0.059	-0.012, 0.144
	FSR vs SC	β_7	0.050	0.049	-0.125, 0.233	0.042	0.042	-0.051, 0.130

*SC as reference

**Rater1, SC as reference

***BCI: Bayesian Credible Interval

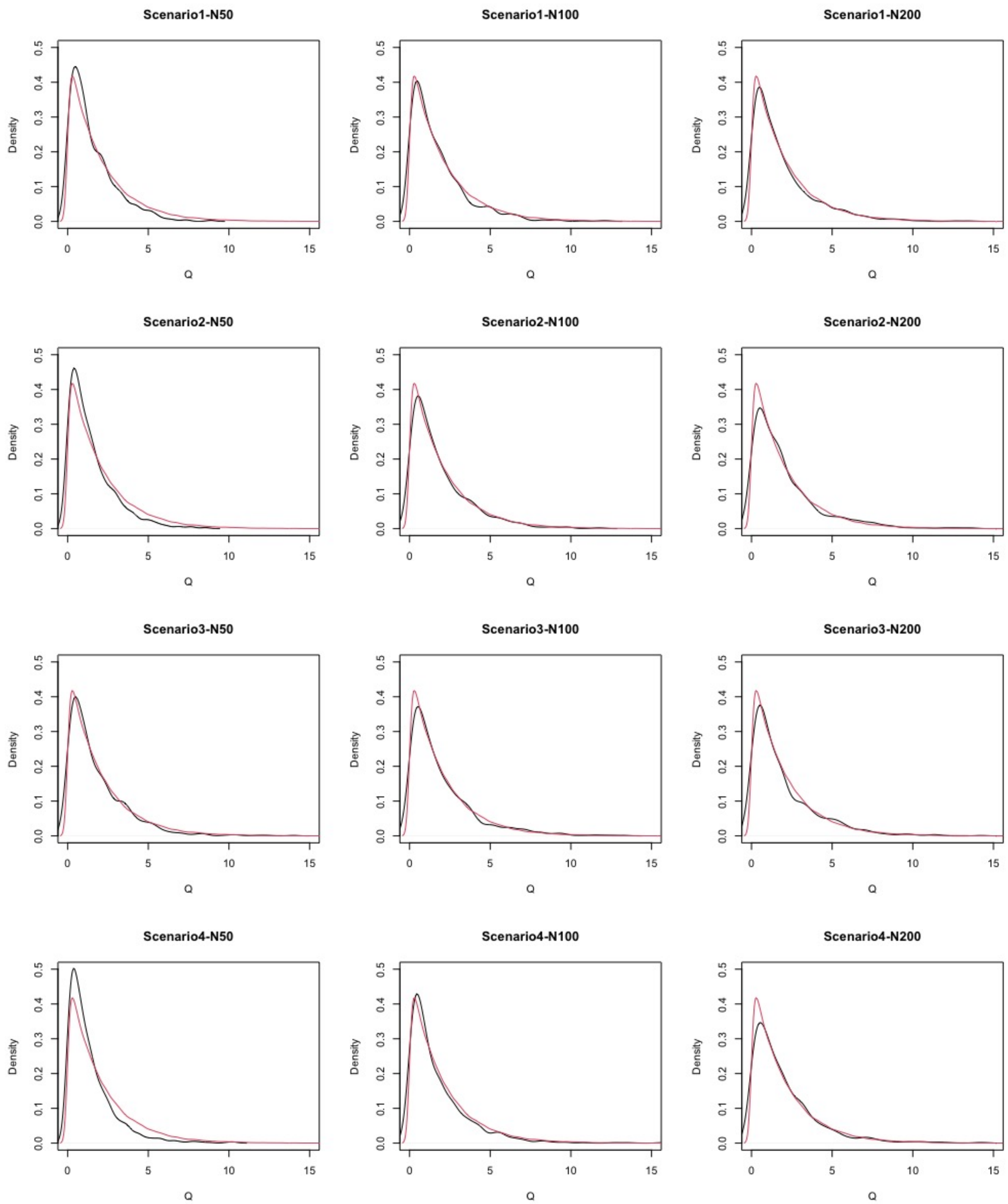

Figure S1. The comparison of Q empirical distribution (black) with chi-square distribution with 2 degree of freedom (red) for scenarios in Table 2.

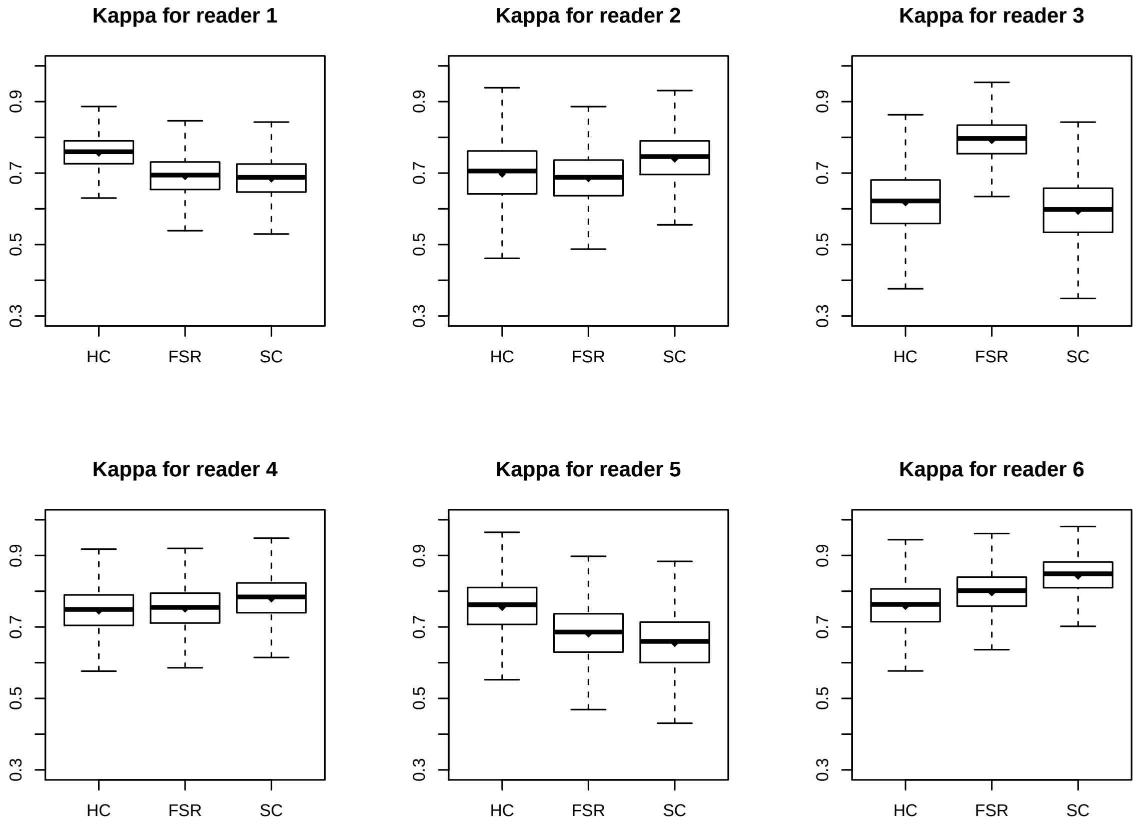

Figure S2. The posterior distribution of kappa of all image formats for all the readers under BMDK

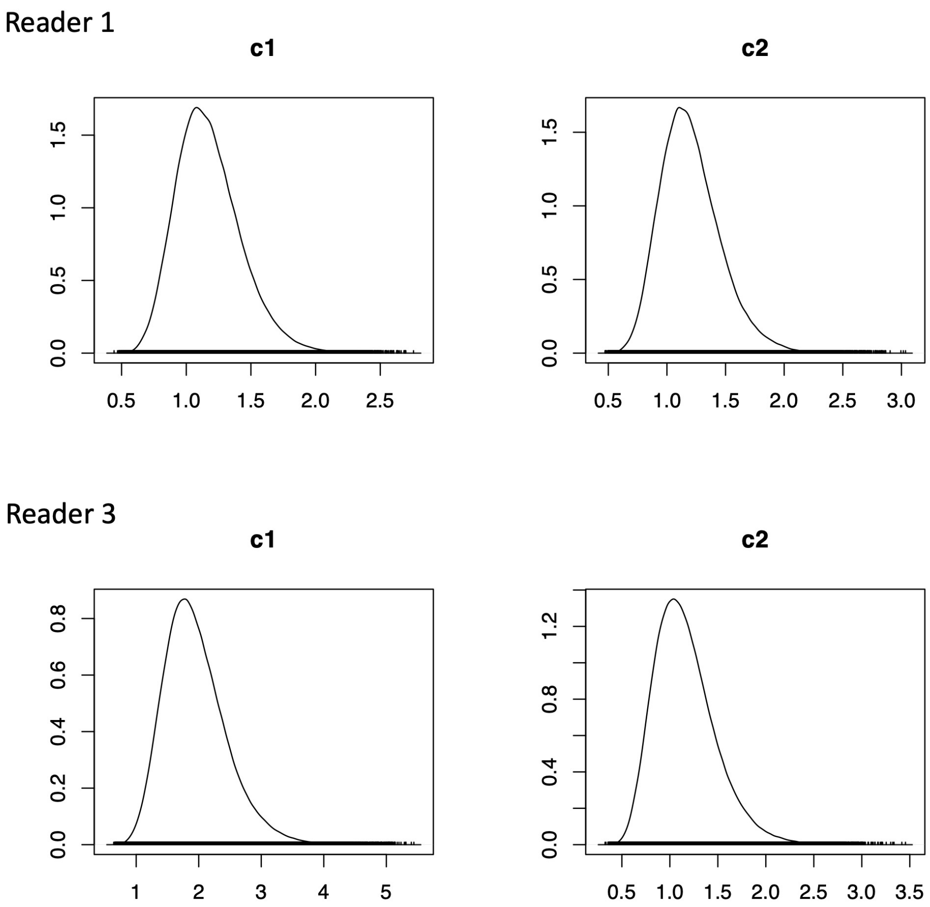

Figure S3. The posterior distribution of C’s for selected readers (BMDK)